DRDC Toronto CR-2004-009

DEVELOPMENT OF MEASURES OF PERFORMANCE FOR EVALUATING THE COMDAT TECHNOLOGY DEMONSTRATOR: POTENTIAL USE OF THE NAVAL COMBAT OPERATOR TRAINER (NCOT) FOR DATA COLLECTION

by:

M.L. Matthews, R.D.G. Webb, A.R. Keeble, T. Lamoureux

Human*systems*, Incorporated
111 Farquhar St., 2nd floor
Guelph, ON N1H 3N4

Project Manager:
Kim Iwasa-Madge
(519) 836 5911

PWGSC Contract No. W7711-017703/001-TOR
Call-Up Numbers 7703-02 & 7703-08

On behalf of
DEPARTMENT OF NATIONAL DEFENCE

as represented by
Defence Research and Development Canada - Toronto
1133 Sheppard Avenue West
North York, Ontario, Canada
M3M 3B9

DRDC Toronto Scientific Authority
Carol McCann
(416) 635-2190

January 2004

Abstract

This report compiles the information contained in four other reports and Technical Memoranda pertaining to the development of Measures of Performance (MOP) for Multi-Sensor Data Fusion technology, and evaluates logistics and the utility of the Naval Combat Operator Trainer (NCOT) to gather these MOPs. Each report/Technical Memorandum is preceded by a summary of its content. This report recounts the development of specific MOPs, the conduct of a proof-of-concept trial at NCOT (with attendant identification of potential system improvements), the outcome of discussions with the contractor responsible for the maintenance and operation of NCOT regarding enhancements, and the investigation of one potential commercial-of-the-shelf solution for some of the improvements described. As a result of identifying these potential improvements, it was recommended that team records created during Navy team training in Operations Room Team Trainer (ORTT) be considered as a better source for deriving MOPs.

Résumé

Le présent rapport contient une compilation de l'information contenue dans quatre autres rapports et documents techniques portant sur l'élaboration de mesures de la performance pour la technologie de fusion de données de multicapteurs (MSDF), ainsi qu'une évaluation de la logistique et de l'utilité du simulateur d'opérateur d'équipement de combat naval (NCOT) en vue de la préparation de ces mesures de la performance. Chaque rapport/document technique est précédé d'un résumé de son contenu. Le présent rapport récapitule l'élaboration de mesures particulières de la performance, la conduite d'un essai de validation de concept au NCOT (l'opérateur identifiant les améliorations possibles à apporter au système), les résultats de discussions menées avec l'entrepreneur chargé de l'entretien et du fonctionnement du NCOT en ce qui concerne les améliorations et l'étude d'une solution commerciale courante possible pour certaines des améliorations décrites. Suite à la détermination des améliorations possibles, il a été recommandé que les dossiers créés par les équipes durant l'instruction d'équipes de la Marine au moyen du simulateur des équipes de salle des opérations (ORTT) soient considérés comme de meilleures sources pour l'établissement des mesures de la performance.

Executive Summary

Over the past several years, Humansystems Incorporated® (HSI®) , under a succession of contracts managed by DRDC Toronto, has identified the critical functions and their associated human behaviours that are the core components of effective Command and Control (C2) performance in the Operations Room of the Halifax Class frigate. The work commenced with a Cognitive Task Analysis (CTA) of the Operations Room Officer (ORO), using a scenario based methodology, to determine the major functions performed by the ORO. The analysis focussed on situation awareness, communication, decision making and workload. Subsequently, with the emergence of the COMDAT TD, HSI was tasked to develop appropriate Measures of Performance (MOPS) based upon the CTA that would be appropriate for assessing the impact of the TD on operational performance. In addition, HSI was requested to develop a Test and Evaluation (T&E) Trial Plan and to evaluate sites and opportunities where appropriate performance data might be collected.

Several independent reports and technical memoranda resulted from this work, and, although complete in themselves, they do not provide a comprehensive view of the work that has been done. Therefore, DRDC-Toronto has requested that the reports be organised into two collections: one based largely on studies relating to the NCOT facility, the other relating to the Operations Room Team Trainer (ORTT).

Five documents are first summarised and then presented in full in this composite report:

The first report: "*Assessing The Impact Of Multi-Sensor Data Fusion On Command And Control Operations In The Halifax Class Frigate: Recommendations For Measures Of Performance And Detailed Test Plan*" (previously published as DRDC CR-2002-018) provides an overall context for the remaining reports. It deals with the development of specific MOPs based upon the prior CTA, provides recommendations for MOPs that may be most useful for assessing the impact of the COMDAT TD and outlines the necessary components of a trial plan to conduct future "human-in-the-loop" (HIL) data collection trials.

The second report: *Technical Memorandum: Findings Of The Test And Evaluation Proof Of Concept Trial At The NCOT Facility (Trial Date March 11-13, 2002) (*previously published as DRDC CR-2002-056) provides details of a an initial assessment of the NCOT facility to support T&E requirements for conducting a "human-in-the-loop" (HIL) trial.

The third report "*Technical Memorandum : Evaluation Of The RGB Spectrum Dgx Digital Graphics Recording System As A Means Of Collecting NCOT Mop Data To Support COMDAT"* also arises from the POC report and assesses the suitability of a COTS product to enhance the existing NCOT recording and playback capabilities. The report recommends the acquisition of such a system should future HIL testing be conducted in NCOT.

The final report "*Technical Memorandum: Discussions With MDA Concerning NCOT Functionality To Support Test And Evaluation Associated With The COMDAT Program"* is a summary of discussions held with MDA (who provide the NCOT software) on changes that would be required in the functionality of the existing software to meet the needs of future T&E trials.

The interested reader is recommended to read this compilation report together with the companion report "Development of Measures of Performance for Evaluating the COMDAT Technology Demonstrator: Potential Use of Training Records from the Operations Room Team Trainer

(ORTT)" in order to be fully informed regarding the development of MOPs and the selection of appropriate venues in which to gather MOPs for the COMDAT TD.

Sommaire

Depuis plusieurs années, Humansystems Incorporated® (HSI®) a, en vertu d'une série de contrats administrés par RDDC Toronto, identifié les fonctions critiques et les comportements humains connexes qui constituent les éléments fondamentaux du commandement et du contrôle (C2) efficaces dans la salle des opérations de la frégate de la classe Halifax. Le travail a commencé par une analyse cognitive des tâches de l'officier de la salle des opérations (OSO), au moyen d'une méthodologie fondée sur divers scénarios, dans le but de déterminer les principales fonctions exécutées par l'OSO. L'analyse a porté principalement sur la connaissance de la situation, la communication, la prise de décisions et la charge de travail. Par la suite, avec l'émergence de la démonstration de technologie d'aide aux décisions de commandement (COMDAT), HSI a été chargée d'élaborer les mesures de la performance appropriées, d'après l'analyse cognitive des tâches convenant à l'évaluation de l'incidence de la démonstration de technologie sur la performance opérationnelle. On a aussi demandé à HSI de préparer un plan d'essai et d'évaluation (E et É) et d'évaluer les emplacements et les occasions où des données appropriées sur la performance pourraient être recueillies.

Plusieurs rapports et documents techniques indépendants ont été produits au terme de ces travaux et, même s'ils sont complets en soi, ils ne donnent pas une vue d'ensemble du travail effectué. C'est pourquoi RDDC Toronto a demandé que les rapports soient regroupés en deux séries : une série fondée principalement sur les études relatives au NCOT et l'autre portant sur le simulateur des équipes de salle des opérations (ORTT).

Cinq documents sont tout d'abord résumés, puis présentés dans leur version intégrale dans le présent rapport global :

Le premier rapport, intitulé « *Assessing The Impact Of Multi-Sensor Data Fusion On Command And Control Operations In The Halifax Class Frigate: Recommendations For Measures Of Performance And Detailed Test Plan* » (publié antérieurement sous le numéro RDDC CR-2002-018), présente le contexte d'ensemble des autres rapports. Il traite de l'élaboration de mesures précises de la performance d'après l'analyse cognitive antérieure des tâches, présente des recommandations concernant les mesures de la performance susceptibles d'être très utiles pour l'évaluation de l'incidence de la démonstration de technologie COMDAT et donne un aperçu des éléments nécessaires d'un plan d'essai en vue de la conduite des futurs essais de collecte de données avec intervention humaine.

Le deuxième rapport, intitulé « *Technical Memorandum: Findings Of The Test And Evaluation Proof Of Concept Trial At The NCOT Facility (Trial Date March 11-13, 2002)* » (publié antérieurement sous le numéro RDDC CR-2002-056), contient des précisions sur une évaluation initiale du NCOT à l'appui des exigences en matière d'essai et d'évaluation pour la conduite d'un essai à intervention humaine.

Le troisième rapport, intitulé « *Technical Memorandum: Evaluation Of The RGB Spectrum Dgx Digital Graphics Recording System As A Means Of Collecting NCOT Mop Data To Support COMDAT* », découle aussi du rapport sur la validation du concept et présente une évaluation de la pertinence d'un produit commercial courant pour l'amélioration des capacités d'enregistrement et de

lecture actuelles du NCOT. Les auteurs du rapport recommandent l'acquisition d'un tel système dans l'éventualité où des essais à intervention humaine soient menés ultérieurement au NCOT.

Le rapport final, intitulé « *Technical Memorandum: Discussions With MDA Concerning NCOT Functionality To Support Test And Evaluation Associated With The COMDAT Program* », est un résumé des discussions tenues avec la société MDA (qui fournit le logiciel du NCOT) sur les changements qu'il faudrait apporter à la fonctionnalié du logiciel pour répondre aux besoins des futurs essais et évaluations.

On recommande au lecteur intéressé de lire le rapport de compilation et le rapport connexe, intitulé « *Development of Measures of Performance for Evaluating the COMDAT Technology Demonstrator: Potential Use of Training Records from the Operations Room Team Trainer (ORTT)* », pour obtenir une information complète sur l'élaboration des mesures de la performance et la sélection des sites appropriés en vue de la préparation des mesures de la performance concernant la démonstration de technologie COMDAT.

THIS PAGE INTENTIONALLY LEFT BLANK

Table of Contents

List of Tables

1. Introduction

Over the past several years, Human*systems*® Incorporated, under a succession of contracts managed by DRDC Toronto, has identified the critical functions and their associated human behaviours that are the core components of effective Command and Control (C2) performance in the Operations Room of the Halifax Class frigate. The work commenced with a Cognitive Task Analysis (CTA) of the Operations Room Officer (ORO), using a scenario based methodology, to determine the major functions performed by the ORO. The analysis focussed on situation awareness, communication, decision making and workload. Subsequently, with the emergence of the COMDAT TD, HSI was tasked to develop appropriate Measures of Performance (MOPS) based upon the CTA that would be appropriate for assessing the impact of the TD on operational performance. In addition, HSI was requested to develop a Test and Evaluation (T&E) Trial Plan and to evaluate sites and opportunities where appropriate performance data might be collected.

Several independent reports and technical memoranda resulted from this work, and, although complete in themselves, they do not provide a comprehensive view of the work that has been done. Therefore, DRDC Toronto has requested that the reports be organised into two collections: one based largely on studies relating to the NCOT facility, the other relating to the Operations Room Team Trainer (ORTT). Thus, in order to obtain a comprehensive and integrated perspective on the work that has been accomplished, the reader is encouraged to review both reports.

THIS PAGE INTENTIONALLY LEFT BLANK

2. Report organization and summary of contents

The main body of this report provides a summary of each of the previously independent reports, followed by each report itself. Where necessary, and for clarity, linking paragraphs between reports have been provided. An overall summary of the reports is provided at the end of the document.

The first report: "*Assessing The Impact Of Multi-Sensor Data Fusion On Command And Control Operations In The Halifax Class Frigate: Recommendations For Measures Of Performance And Detailed Test Plan*" (previously published as DRDC CR-2002-018) provides an overall context for the remaining reports. It deals with the development of specific MOPs based upon the prior CTA, provides recommendations for MOPs that may be most useful for assessing the impact of the COMDAT TD and outlines the necessary components of a trial plan to conduct future "human-in-the-loop" (HIL) data collection trials.

The second report: *Technical Memorandum: Findings Of The Test And Evaluation Proof Of Concept Trial At The NCOT Facility (Trial Date March 11-13, 2002)* (previously published as DRDC CR-2002-056) provides details of a an initial assessment of the NCOT facility to support T&E requirements for conducting a "human-in-the-loop" (HIL) trial.

The third report "*Technical Memorandum: Evaluation Of The RGB Spectrum Dgx Digital Graphics Recording System As A Means Of Collecting NCOT MOP Data To Support COMDAT"* also arises from the POC report and assesses the suitability of a COTS product to enhance the existing NCOT recording and playback capabilities. The report recommends the acquisition of such a system should future HIL testing be conducted in NCOT.

The final report "*Technical Memorandum: Discussions With MDA Concerning NCOT Functionality To Support Test And Evaluation Associated With The COMDAT Program"* is a summary of discussions held with MDA (who provide the NCOT software) on changes that would be required in the functionality of the existing software to meet the needs of future T&E trials.

THIS PAGE INTENTIONALLY LEFT BLANK

3. Report:

ASSESSING THE IMPACT OF MULTI-SENSOR DATA FUSION ON COMMAND AND CONTROL OPERATIONS IN THE HALIFAX-CLASS FRIGATE: RECOMMENDATIONS FOR MEASURES OF PERFORMANCE AND DETAILED TEST PLAN.

3.1 Summary

3.1.1 Background

This report was the first in the COMDAT human engineering series. The work reported was commissioned to develop potential Measures of Performance (MOPs) for Halifax (HFX)-Class operations (Ops) room teams. In particular, these MOPs needed to be practical from the point of view of data collection in a simulated Ops room team environment, such as the Naval Combat Operator Trainer (NCOT) and the Ops Room Team Trainer (ORTT). The MOPs also needed to address the range of cognitive demands (e.g. Situation Awareness, communications, decision-making) faced at both the individual level and the team level.

Having developed MOPs, this work was also to develop a detailed trial plan for MOP data collection in a Test and Evaluation (T&E) environment. This data collection would assess the adequacy of the MOPs in terms of what they reveal about the Ops room team and its individual operators, as well as the adequacy of the simulated environment for conducting such data collection exercises.

3.1.2 Findings

3.1.2.1 Development of Measures of Performance

The development of MOPs was based on a body of work already completed. This work included a detailed Cognitive Task Analysis (CTA) of the ORO, a literature review of MOPs for Command and Control (C2) in general and Navy C2 in particular, and the COMDAT 1 Technology Demonstrator Program (TDP) system design manual. The approach taken was to define the ORO functions impacted by MSDF and, their criticalities and frequencies. The impact of MSDF, the criticality of the functions and the frequency of the functions were all assess on a three-point scale (high-medium-low).

T&E TASK#	ORO Function/Task from CTA Validation Function Flow	COMDAT1 Reference (as above)	Focus of MOP from Framework for Evaluation	Priority
	2.1 BUILD MAINTAIN GLOBAL PICTURE			
1	Check Mission picture - 2.1.1	1,2,3	SA1, COM	1
2	Relate new info to Ops Room picture - 2.1.2	1,2,3	SA2,3 COM DM	1
NA	Relate new info to Ship Picture - 2.1.3	NA	-	-
NA	Relate new info to Task Group Picture - 2.1.4	NA	-	-
3	Relate new info to Recognised Maritime Picture (formerly MTP)	1,2,3	SA2,3	1
4	Relate to Recognised Air Picture -2.1.5	1		1
5	Relate to Maritime Surface Picture and Sub-surface Picture -2.1.6	1,3		1
6	Relate to Wide Area Picture*	2,3		3
7	Assess threats*	1,2,3	Implicit in SA2	1
8	Switch attention between different pictures**	2,3	SA2,3	3
	2.4 MANAGE SHIP			
9	Assess sensors - 2.4.1.2	1,2,3	SA1 DM	1
NA	Assess weapons - 2.4.1.3	NA	-	-
10	Manage ship surveillance - 2.4.2	1,2,3	SA1,2,3, DM, COM	2
4***	Ensure contact classified – 2.4.3.1	1,2,3	SA1,2,,3 DM, COM	1

Table 3.1: Mapping of COMDAT1 Technology onto ORO functions, evaluation focus and priorities for MOP development.

In total, nine ORO CTA derived functions were selected (see Table 3.1) for this analysis. The single function "Relate (new info) to RAP and RMP" was further expanded to include the following discrete functions:

- Relate new info to Recognised Maritime Picture (formerly MTP);
- Relate to Recognised Air Picture;
- Relate to Maritime Surface Picture and Sub-Surface Picture.

Additionally, three functions were added: Assess Threats, Relate to Wide Area picture and Switch Attention between Different Pictures.

The impact of MSDF on some functions was undetermined, and in other cases their criticality and frequency were ambiguous. In such instances, the functions were bookmarked in case there was a requirement to pursue them at some later time.

Since the behaviours and their associated MOPs dealt with all ORO functions, the next step was to identify the subset of MOPs that were most relevant to the focus areas of the COMDAT 1 TDP and to map these on to the domain of measurement (Situation Awareness levels 1: perception, 2:

comprehension and 3: projection, Communications, Decision Making). Finally, the MOPs were prioritised according to the following 4 criteria:

- The importance and criticality of the function and the degree of MSDF impact;
- The likelihood of a function being implemented earlier than others;
- The likely difficulty of implementing a function;
- The likely difficulty of gathering the MOPs associated with a function.

In total, 67 detailed MOPS were established relating to a broad range of Operations Room C2 functions..

3.1.2.2 Test and Evaluation Plan

An incremental approach to testing was chosen due to a number of factors: the initial focus on Above-Water Warfare (AWW); uncertainty about how the TDP would impact the Wide Area Picture (WAP) and surface/subsurface integration; the need to gain experience in developing scenarios in NCOT and ORTT; and uncertainties around the logistics of Human-In-the-Loop (HIL) testing. In order to determine the most appropriate test and evaluation environment for collecting MOP data, it was recommended that further investigation of ORTT be made alongside the investigation of NCOT.

The following main considerations for conducting a T&E trial were described in detail.

- Logistical requirements;
- Scenario development;
- Detailed test plan;
- Proof of concept;
- Pilot study;
- Main study;
- Data analysis;
- Conclusions and lessons learned.

A number of issues with respect to data reliability were also discussed, including variability between ORO subjects, ORO workload variance, tradeoffs due to the research design, sample size effects, realism, generalisability, the implementation of MSDF technology into the simulator and potential problems with getting access to a sufficiently large number of test participants to ensure valid and reliable data.

3.1.3 Conclusions

The initial recommendation was that the initial T&E trials be conducted in the NCOT facility, which appeared to provide all of the requirements for the first stage of evaluating the most immediately emerging TDP functionality, while minimising the logistical overhead to support the trial.

The report concluded with a number of relevant references, a description of the Detect-to-Resolve cycle (the Ops Room function that appears to be most obviously impacted by the COMDAT TD) and an initial follow-up evaluation of the ability of NCOT to support the gathering of MOPs.

3.2 Introduction

This report represents a continuation of work aimed at improving decision support to the Operations Room (Ops Room) functions of the Navy Halifax class ships. The present work arises out of a contract from DRDC to Human*systems* Inc. to develop a comprehensive Test and Evaluation (T&E) program to measure Ops Room functions and assess any future impact by MSDF technology.

This technical memorandum addresses specific Statement of Work items:

- 3.1.2 Develop Potential Measures of Performance (MOPs)
- 3.1.6 Create plan for piloting data collection
- 3.3.4 Create detailed evaluation report.

In this section the general approach and strategy to developing MOPs are described and issues concerning terminology and usage of the term "pictures" are addressed.

3.2.1 Developing the MOPs

The approach to developing the MOPs has been guided by a number of relevant sources that include:

- The ORO cognitive task and function flow analyses (references 1,2)
- A brief review of the Navy C2 literature concerning concepts of MOPs
- Previous analysis of methods for evaluating and measuring C2 system performance (reference 3)
- The COMDAT1 Technical Demonstrator (TDP) system design (references 4,5).

3.2.1.1 The ORO Cognitive Task Analysis

The foundations for the current work are based upon a cognitive task analysis (CTA) conducted with a representative cross section of serving Operations Room Officers (OROs) (reference 1). A scenario based walkthrough of a hypothetical mission, somewhat similar to Op Crater[1], was used to elicit the major operational goals that were required to achieve mission objectives. The information, situation awareness and communication requirements to achieve these goals were also

[1] A scenario used by the Navy in Ops Room team training

identified. Initially goals were grouped into three categories: Coming on Watch, Foreground and Background goals. Subsequently, a validation of the initial findings was conducted on a different group of OROs (reference 2). As a result of this, Foreground and Background Goals were revised to become Threat-Related Goals and Ops Room Management goals. In addition, function flow analyses were created based upon the goal structure and SME input.

Using these analyses, the critical, core functions and tasks performed by the ORO to support generic mission goals were identified. These then served as the basis for focussing current efforts in determining appropriate MOPs that could be used to measure ORO/System performance in these areas.

3.2.1.2 Existing MOPs in the Literature

A brief literature review was conducted of military and particularly Navy MOPs that would be directly relevant to the current task. The focus of the review was on process-related measures rather than outcome measures of effectiveness (e.g. number of targets destroyed, number of casualties taken). In general, nothing specific was found to be applicable, other than recommendations to use accuracy/error and time on task measures, supplemented by subjective ratings where appropriate. Therefore, previous work on the measurement of C2 effectiveness was used to guide the selection of MOPs with a focus on measures of situation awareness, communication and decision making (reference 3).

3.2.1.3 Previous Analysis of Methods for Evaluating C2 Systems

A comprehensive review has been conducted on a variety of methods and measures that have been discussed or implemented in the search for appropriate ways to assess C2 performance in a broad spectrum of military contexts (reference 3). In general, the present approach follows the recommendations of this report by focussing largely on measures of performance that assess the underlying human-system processes that determine operational outcomes. The three process areas that are the focus of attention in the present analysis are: Situation Awareness, Communication and Decision Making. Specific methods and approaches for gathering performance data in these areas are provided in this report.

3.2.1.4 The COMDAT1 Technical Demonstrator (TDP) System Design

Another important element in focussing the scope of the *initial* set of MOPs has been the COMDAT1 TDP. At present, the TDP in many areas of MSDF[2] appears to be a concept in technical evolution, at least as far as we have been able to ascertain from available documentation and discussions with the Scientific Authority. Thus, there remains some uncertainty as to how and what functions will be implemented at the level of the ORO and other members of the Ops Room teams. Thus, while the focus of the CTA and the present mandate are the functions of the ORO, it would appear that in the initial stages the TDP effort may have more impact upon tasks performed by Ops Room teams in the different warfare areas.

Notwithstanding this uncertainty, the strategy that we have adopted in developing the MOPs is to match the ORO CTA goals with the COMDAT1 system concept as it stood at the time of writing. This allows the ORO functions to be identified that will most likely be impacted by MSDF

[2] We use the definition of MSDF as outlined in reference 4, multi-*source* data fusion (i.e. not just *sensor*)

concepts embodied in the initial TDP. From these functions we can then look at the specific tasks with a view to developing the appropriate MOPs.

The review of the available literature and technical memoranda and the information gathered from technical briefings indicate that, in the first instance, the TDP will focus on the integration of air sensor information at the track level, to aid tasks related to track detection, maintenance and possibly classification. As noted in previous discussions with the relevant Scientific Authorities, and based upon the CTA, normally the ORO has little direct involvement in the processing of information at this level or in the direct conduct of these tasks. Rather, the ORO's interest is managing processes performed by various members of the Ops Room team. The ORO will also under some circumstances share some of the air warfare duties with the SWC when the workload exceeds the capability of the SWC to cope. In the role of manager of the teams, the ORO has a major interest in the product of the various teams' activities in terms of building his tactical situation awareness, assessing and evaluating the tactical situation, planning actions, monitoring responses and co-ordinating the Ops Room surveillance function. Other aspects of the TDP, as outlined in the system concept (reference 5, figure 1) appear to have a more direct impact on some of the ORO core functions. These include the integration of the wide area picture, tactical picture, and the above and underwater pictures. This concept diagram also implies the integration of *non-sensor* sources of information. Exactly what technology is involved and how it will be implemented remains unclear at the moment, therefore we have made our best guess about the potential impact upon the ORO and have developed MOPs accordingly.

Further into the future, beyond COMDAT1, there remains the possibility that information fusion may impact tasks such as managing various aspects of the Ops Room capability (process refinement in the JDL model), which are core ORO functions. For now, we have put these issues on the back burner, given the immediate priorities of assessing the potential impact of MSDF in the COMDAT1 TDP.

Before proceeding with the description of the MOPs, we believe it would be useful at this point to clarify some terminology issues concerning the use of the word "pictures" to avoid subsequent potential confusion in interpretation.

3.2.2 Terminology Concerning "Pictures"

With respect to current Navy C2 thinking[3], the term Maritime Tactical Picture (which was used extensively in the CTA and reflected existing Navy terminology) appear to be have been replaced by the term Recognised Maritime Picture. The latter is part of a hierarchy of "pictures" as outlined below.

Common Operational Picture (COP)

Recognised Maritime Picture (RMP)

Maritime Surface Picture (MSP)*

Maritime Sub-surface Picture (MSubP)*

Recognised Air Picture (RAP)

Recognised Land Picture (RLP)

[3] Based on references 6 and 7 and information provided by the Scientific Authority.

*Note: the above abbreviations represent labels of convenience for this report and do not necessarily reflect Navy terminology.

The *Common Operational Picture* is the composite of all operational information provided by each element (Maritime, Air, and Land)[4], fused and filtered as necessary to support the requirements of each specific user. The subordinate Recognised Pictures represent the operational information generated within each element, and specifically tailored for use by that element. For example, in addition to data on ships and submarines, a Maritime Commander's picture includes data on air forces (ASW helicopters, Maritime Patrol Aircraft, tactical fighters in combat air patrols over ships, hostile strike aircraft, missiles, etc). Data on ships is also of interest to an Air Commander operating near maritime areas. A Land Commander operating near coastlines will also be interested in data pertaining to naval and air forces in his area. There could also be other Recognised Pictures in addition to the above, and other sub-sets within each recognised picture. The underlying concept is that each picture represents the view of interest to a specific Command level; the COP represents the composite of all available data.

Formally, the RMP has been defined as "….a compilation of all source data relating to a specific ocean area, known at a given time and disseminated following evaluation and validation….". The term "recognised" means that the plot is an evaluated and validated interpretation of available information…". This can result in a recognised contact remaining "unknown" after the evaluation process.

Based upon the above, it would appear that the ORO will be dealing with pictures below the level of the COP, for the most part. The ORO's primary interest and responsibilities will be at the RMP level with a frequent requirement to interact with pictures below this level. Presumably, if TG duties included responsibilities for air warfare, then the ORO concerned would also have a primary involvement at the RAP level.

These definitions and concepts of the various categories of pictures will be subsequently used in this report as we identify the areas for specific MOPs.

3.3 ORO Functions Impacted by MSDF in COMDAT1

In this section we describe how we used the functions resulting from the CTA and CTA validation as a starting point for determining areas for which MOPs should be developed. From this initial set of functions, individual functions were either eliminated or augmented. Reasons for elimination included low relevance to MSDF improvements, low criticality or frequency and logistical considerations with respect to implementation for T&E. The list was augmented to include more detailed functions in specific areas that were to be the immediate focus of the COMDAT1 TDP.

3.3.1 Function Selection and Filtering

The first step in the process of determining MOPs was to identify the major ORO goals and functions from the CTA validation study that could be affected by MSDF technology, considered in its broadest form. These functions are shown in Table 3.2 and are taken directly from the function flow diagrams using the same reference numbers (reference 2 Annex A). Each function was then rated by the HSI team (which included one experienced ORO) in terms of operational criticality and frequency and the likely impact of MSDF technology. Then, in order to focus the

---

[4] This includes the wide area picture (WAP), local area picture (LAP) and organic and non-organic sources.

MOP development effort, functions were eliminated from this list if they met any of the following criteria:

1. Low impact from MSDF technology (at least, as presently conceived in COMDAT)
2. Tasks done off line from the CCS
3. All tasks involving preparation and planning, which are considered to be out of scope by the Scientific Authority with the present focus on operational performance.

Functions eliminated by this process are shown with a grey background in Table 3.2.

Among the items eliminated by this process are a number of critical, **process management** functions that are core to the OROs overall operational focus. These have been temporarily set aside for now, for two reasons. First, we believe them to be generally beyond the *immediate* impact of the COMDAT1 TDP. Second, we have not seen any specific conceptualisations of the kind of COMDAT technology that could support ORO decision making in the conduct of these functions. The functions temporarily set aside are:

- Assess teams (2.4.1.4)
- Assess comms (2.4.1.5)
- Assess schedule (2.4.1.6)
- Optimise capability (2.4.1.7)
- Plan ship response (2.4.3.2)
- Implement response (2.4.3.3)
- Assess ship response (2.4.3.4)

If future TDPs incorporate some of the ongoing R&D concepts from DREV concerning situation analysis, impact assessment and process refinement, then the development of a wider range of MOPs for functions related to monitoring and managing processes and people will be pursued.

ORO Function/Task from CTA Goal List or Function Flow (ref.1)	MSDF IMPACT	CRITICALITY	3.3.1.1.1 FRE QUENCY
1. Prepare for watch			
Review external environment - 1.1.1	LOW	MED	LOW
Update awareness - 1.1	LOW	HIGH	LOW
Review schedules - 1.1.2:	LOW	HIGH	LOW
Review capability - 1.1.3	MED	HIGH	LOW
Review ongoing Ops Room team tasks (2-5 mins) - 1.1.4	LOW	HIGH	LOW
Confirm authority - 1.1.5	LOW	HIGH	LOW
Review Mission Picture - 1.1.6	HIGH	HIGH	LOW
Comprehend briefing from outgoing ORO - 1.1.8	LOW	HIGH	LOW
Modify pre-plan/procedures for watch - 1.1.9	LOW	HIGH	LOW
Prepare workstation 1.2	LOW	HIGH	LOW
Focus OR team on issues for coming watch. (2 mins) - 1.3	LOW	MED	LOW
2. CONDUCT WATCH:			
Build /Maintain Global Picture 2.1 (see items below)			
Check Mission Picture- 2.1.1	HIGH	HIGH	HIGH
Relate (new info) to Ops Room picture - 2.1.2 (includes helo, weapons, CCS/SSD, organic sensors, comms.)	HIGH	HIGH	HIGH
Relate (new info) to Ship Picture - 2.1.3	MED	MED-HIGH	MED
Relate (new info) to Task Group Picture - 2.1.4	MED	HIGH	MED
Relate (new info) to RAP 2.1.5 and RMP 2.1.6 (formerly referred to as MTP)	HIGH	HIGH	HIGH
Relate (new info) to environment picture - 2.1.7	ORO: LOW Operators: HIGH	HIGH	MED
Relate new info to ship schedule - 2.1.8	LOW	HIGH	MED
Prioritise mission needs -2.1.9	HIGH	HIGH	MED
Check own authority to act - 2.1.10	LOW	HIGH	MED
Plan watches ahead 2.2	LOW	MED	LOW
Manage Own Information Exchange - 2.3	LOW	HIGH	HIGH
2.4 MANAGE SHIP (1)			
Manage Ship Capability -2.4.1 (see items below)			
Assess sensors - 2.4.1.2	HIGH	HIGH	MED
Assess weapons - 2.4.1.3	HIGH	HIGH	MED
Assess teams - 2.4.1.4	LOW	HIGH	MED
Assess communications – 2.4.1.5	HIGH	HIGH	HIGH
Assess schedule - 2.4.1.6	LOW	HIGH	MED
Optimise capability - 2.4.1.7	HIGH	HIGH	MED
Manage ship surveillance - 2.4.2	HIGH	HIGH	HIGH
Manage ship response (see items below) - 2.4.3			
Ensure contact classified – 2.4.3.1	HIGH	HIGH	HIGH
Plan ship response - 2.4.3.2	MED	HIGH	HIGH
Implement response - 2.4.3.3.	MED	HIGH	HIGH
Assess ship response 2.4.3.4 (see below)			
Determine outcome for contact of interest 2.4.3.4.1	HIGH	HIGH	HIGH
Determine outcome for own ship - 2.4.3.4.3	HIGH	HIGH	HIGH
Report- 2.4.3.5	MED	HIGH	HIGH

Table 3.2: Functions derived from the CTA and CTA Validation rated in terms of mission criticality and frequency and potential MSDF impact.

Notes:

(1) This analysis is based upon a threat-related mission.

We now offer a few points of commentary on two of these decisions.

Optimise capability may have a high MSDF impact, however, the actual role MSDF may play in facilitating the performance of this function is not clear. It is possible that the impact of MSDF will be more in the form of assessing capability rather than assisting the process of rectifying any capability shortcomings. The assessment of capability can be thought of as having two parts: the assessment of sensors and the assessment of the deployment of appropriate personnel for the tasks at hand. There appears to be a clear role for lower MSDF levels for the former, but the latter would involve MSDF level 4 processes refinement, for which there appear to be no specific technological solutions on the immediate horizon. Thus, we have not developed specific MOPs for this function right now, but will use the MOPs associated with the function "assess sensors" to address the issue of assessing capability, where appropriate.

Determine outcome for COI. Clearly, this function has possible MSDF implications for the improved recognition of the changed status of hostile air tracks that might result from an engagement. As such it might be seen to have a high priority for MOP development. However, our reasons for setting it aside for now are largely practical, since it will demand a greater level of complexity in the running T&E trials, than the payoff may warrant. The additional Ops Room team and ORO functions that would be required to simulate the engagement of a threat presents a major additional overhead when considering the infrastructure and logistical requirements for the T&E scenario. Further, we believe that the specific changes in air tracks that arise from an engagement outcome can be simulated using the proposed threat events that are listed below.

Finally, the Scientific Authority may wish to evaluate priorities for the future development of MOPs in core areas of ORO functioning for some of those functions currently filtered out by the above process, even in the absence of conceptions of how technology could be of assistance. Several of these functions relate to the ORO's core role as the manager of processes and people and ultimately determine the overall *effectiveness* of the Ops Room.

3.3.2 Function Augmentation

Before proceeding further with refining the list, we saw a need to expand the depth of analysis for some the detailed Ops Room tasks that might be impacted by the more immediate implementation of COMDAT1 technology. This was considered necessary in view of the proposed focus of the application areas for the COMDAT1 TDP and some uncertainty as to whether the impact of the technology would fall upon the ORO or other members of the Ops Room teams. This analysis comprised a more detailed review of Ops Room processes with a Navy SME in order to gain a better understanding of where the COMDAT technology would impact. One result of this was the development of process flow diagrams that identified the major Ops Room tasks in the detect-to-resolve process that appears to be major area affected by the COMDAT1 TDP (see Annex A).

As a result of this process and insights derived from analysis of the JDL MSDF concept, it became evident that some revisions would be appropriate to the original CTA decomposition to accommodate this new information obtained. Specifically, the CTA and function flow analysis in the area of situation and threat assessment needed to be updated and enhanced.

In the CTA and subsequent function flow analysis, situation assessment (and the associated sub-functions relating to maintaining the air, surface and sub-surface pictures) were seen to be implied within the overall function of "*Build and maintain global picture"*. In retrospect, it would appear that the verbs "build" and "maintain" overlook the need for "comprehension" and "understanding" in terms of tactical relevance that are the very reasons for building and maintaining the picture.

Given the importance of comprehension of the built picture to the assessment of the tactical situation, we propose that situation assessment should become an explicit focus for MOPs.

Situation assessment may be seen as a generic term that could be used across a variety of combat related or peacetime scenarios. However, in the context of typical training scenarios, situation assessment largely involves a narrower set of goals dealing with the evaluation of *threats*. We will therefore adopt this more restrictive usage as the basis for conceptualising MOPs, since the present focus of MSDF technology appears to be on improving information to the Ops Room team concerning contacts that are specifically threat related.

In terms of the existing function flow analysis, the "*assess situation*" function should probably be inserted as an explicit step after functions 2.1.2-2.1.8 (relating information to various pictures) and before 2.1.9 *Prioritise mission needs* and 2.1.10 *Check own authority to act*. Thus, in the tables and sections that follow, MOPs for threat assessment will be provided even though this function does not map directly onto the organisation of the functions provided by the CTA or function flow analysis.

A second area that we believe to be of importance that was captured in the original CTA, but not reflected specifically in the function flow diagrams, concerns the issue of *attention switching by the ORO*. This was identified as a critical ORO function with respect to both the evaluation of the tactical picture across all warfare domains as well as in managing and monitoring Ops Room processes and personnel. Therefore, we have included attention switching as a potential area for MOP development.

A third area of potential focus for MOPs resulting from a review of the COMDAT1 TDP system concept concerns the integration of wide area picture (WAP) information. This was not explicitly differentiated from the Air, Surface or Sub-surface pictures in the CTA or subsequent validation. However, given that the TDP will be designed to facilitate the integration of the wide and local area pictures, no matter the domain, it would be prudent to consider this as an area for performance assessment and hence MOP development.

To summarise, the above list of functions from the CTA validation has been augmented in the following three areas:

- Threat assessment
- Attention switching
- Integration of information from WAP

We now return to the issue of MOP development.

3.3.3 Priorities for MOP Development

Table 3.3 shows the list of functions resulting from the initial filtering and augmentation process outlined in 2.1 and 2.2. In order to establish priorities for the initial development of MOPs for these functions, it has been necessary to make our best guess as to what MSDF technology is likely to emerge in the immediate future. In the absence of any other concrete conceptualisation, we have used the DREA System Concept (references 4,5) for guidance, which outlines the following areas that will be the focus of COMDAT1. The ordering below is based upon information that has been provided by the Scientific Authority on the planned development and implementation sequence for the TDP.

1. AWW tactical picture: implications for RAP, MSP
2. Integration of AWW/WAP: implications for RAP, MSP, RMP

3. Integration of AWW, UWW and WAP: implications for RAP, MSP, MSubP, RMP.

Before relating these areas of focus of the TDP to the Ops Room Functions, some elements of uncertainty surrounding the above should be noted. With respect to step 1, the documentation that has been made available to us, and information communicated from the Scientific Authority concerning the trials that have been conducted, suggest that the TDP emphasis has been on the fusion of *all air sensor and Link 11 data*. This would be of primary relevance to AWW. We have found no mention of the specific Navy systems that would form the basis for fusion of data, or integration of information that would address the other component of AWW, namely the MSP. A further consideration that requires clarification concerns the integration of the UWW sensor information in step 3. There are at least two possibilities. First, by analogy with the integration of air sensor data, the UWW sensor data (CANTASS, HMS, SPS) could itself be integrated to provide the best available UWW picture. Or, second, the existing streams of already processed information from these three systems could simply be provided at the track level and displayed upon a common integrated display that would also include air and surface tracks. Clarification of these areas will be necessary at some point in time to ensure we fully understand which Ops Room functions and personnel will be impacted by the various possible configurations of data and information elements.

Notwithstanding this ambiguity and uncertainty, our next step has been to map the three TDP areas against each of the remaining subset of ORO functions in terms of where the TDP is likely to have impact. The outcome of this is shown in column three of Table 3.3. Where we could not see an obvious impact of the COMDAT1 technology, we have indicated not applicable (NA). We have also indicated the appropriate domain of measurement[5] in column four, where SA= situation awareness (levels 1,2,3[6]), COM= communication (that is all aspects of information exchange, not simply those aspects using only audio comm nets) and DM= decision making.

The final column of Table 3.3 provides a recommendation of priorities for developing the MOPs. These recommendations are based upon a consideration of four factors. First, the importance and criticality of the function and the degree of MSDF impact. Second, our current perception of the timeline for the TDP development - functions that are likely to be impacted earlier are assigned higher priority. Third, our intuition about the difficulty in implementation of some of the associated, core technologies and the clarity of the concepts. Fourth, the logistical difficulty in collecting MOP data - functions that will require complex scenarios across several warfare areas and involving several Ops Room teams are assigned lower priority.

It should be noted that in this table we have changed the function 2.1.6 *Relate to MTP to Relate to RMP* to be consistent with more current Navy terminology. We have also further decomposed this into the constituent elements of "Relate to… Maritime Surface Picture (MSP)",…. Maritime Sub-Surface Picture (MSubP) in order to provide a better mapping onto the separate functional areas of the COMDAT1 TDP. We have also included the WAP in the consideration for developing MOPs, even though this was not identified as such in the CTA, since this is clearly a focus of the MSDF technology and can be considered as contributing to the RMP.

5 Based upon reference 3

6 SA1=detection of new information; SA2=integration and comprehension of information; SA3=projection of future state.

T&E TAS K#	ORO Function/Task from CTA Validation Function Flow	COMDAT1 Reference (as above)	Focus of MOP from Framework for Evaluation	Priority
	2.1 BUILD MAINTAIN GLOBAL PICTURE			
1	Check Mission picture - 2.1.1	1,2,3	SA1, COM	1
2	Relate new info to Ops Room picture - 2.1.2	1,2,3	SA2,3 COM DM	1
NA	Relate new info to Ship Picture - 2.1.3	NA	-	-
NA	Relate new info to Task Group Picture - 2.1.4	NA	-	-
3	Relate new info to Recognised Maritime Picture (formerly MTP)	1,2,3	SA2,3	1
4	Relate to Recognised Air Picture -2.1.5	1		1
5	Relate to Maritime Surface Picture and Sub-surface Picture -2.1.6	1,3		1
6	Relate to Wide Area Picture*	2,3		3
7	Assess threats*	1,2,3	Implicit in SA2	1
8	Switch attention between different pictures**	2,3	SA2,3	3
	2.4 MANAGE SHIP			
9	Assess sensors - 2.4.1.2	1,2,3	SA1 DM	1
NA	Assess weapons - 2.4.1.3	NA	-	-
10	Manage ship surveillance - 2.4.2	1,2,3	SA1,2,3, DM, COM	2
4***	Ensure contact classified – 2.4.3.1	1,2,3	SA1,2,,3 DM, COM	1

Table 3.3: Mapping of COMDAT1 Technology onto ORO functions, evaluation focus and priorities for MOP development.

Notes on Table

* No specific counterpart in CTA function analysis
** Taken from CTA, this not explicitly identified as a unique function in the function flow analysis
*** To be measured as part of task 4.

3.3.3.1 Rationale for Final Selection of Tasks to be Measured

A number of the functions in Table 3.3 have been assigned lower priority for measuring, despite having high criticality and medium or high MSDF impact. This section provides the rationale for those assignments.

Switch attention between different pictures is an ORO process identified in the CTA that falls within the "Build and Maintain Global Picture" function but could not identified as a unique sub-function in its own right in the function flow diagrams. While we believe this process to be one of the most critical components for effective ORO performance, the rating of 3 has been assigned because of logistical considerations in its measurement. In order to elicit the behaviour of switching between warfare domains, a greater complexity of scenario will be required to include events of interest across all warfare domains. We believe that we will be in a better position to formulate MOPs for this function, once some experience has been gained in scenario development, data

collection and proof of concept testing for MOPs using less complex scenarios. Lessons learned should then enable us to rapidly formulate the most appropriate MOPs for this function.

Relate information relevant to the WAP would seem to be directly impacted by the intention in COMDAT1 to integrate the WAP and LAP pictures, by making available to the ORO information from the GCCS/ MCOIN systems. However, we have not developed specific MOPs for this function because at present the concept of how this is to be realised seems somewhat unclear. For example, the GCCS currently provides *information* (that is processed tracks) rather than *data* (radar returns). Hence, it is not clear how such distal information can be "fused" with local radar data - since they are addressing essentially different subsets of contacts at different ranges. Our current belief is that the proposed approach in the TDP for the provision of the WAP to the ORO appears to be more in the format of *information aggregation* rather than data fusion. One design approach that is being considered to implement this concept is to provide dual displays - with the lower providing information on the LAP and the upper the WAP. That is, information at the track level for the wide and local pictures will be co-located. This approach is somewhat different from concepts that involve *fusion* or *integratation* at the semantic level. This design solution would improve upon the current situation whereby the LAP and WAP are displayed on monitors that are in different places in the Ops Room. However, it remains to be seen whether this relocation of displayed data results in providing the ORO with a seamless view of the local and long-range tactical pictures, as described by the "fused scene" concept in reference 6. Notwithstanding this uncertainty concerning the level at which data or information are to be integrated and the exact design solution that will be adopted, it will probably be the case that MOPs developed and tested for somewhat similar functions relating to the detection, integration and interpretation of new information for the RAP or MSP or MSubP may be readily modified to assess integration of the LAP and WAP.

Manage ship surveillance is also rated HIGH for MSDF impact, where we believe that the contribution of the technology is likely to be in the area of providing better information to the ORO concerning the performance of sensors and integration of information from the WAP and across warfare domains. Many of these capabilities can be assessed by MOPs outlined elsewhere for related functions. The other aspect of managing surveillance concerns the deployment and monitoring of personnel performing surveillance functions and the development of surveillance plans in light of changing circumstances. For the former, a preliminary set of MOPs is suggested relating to the ORO's ability to observe when there are errors or problems in the standard detect-to-resolve process (in the specific case of air threats).

3.3.4 Recommended MOPs

The initial focus of MOP development is on the functions to which we have assigned a priority rating of 1 or 2 in Table 3.3. In specifying the MOPs in the table that follows, some of the above functions have been grouped together because of overlapping MOPs. For example relating new information to the RMP is highly related to the function of building awareness across domains and forming an integrated tactical picture. In this initial set of MOPs, greater attention has been placed upon the function "Relating new information to the RAP" since integration of air sensor data will be the initial focus of the COMDAT1 TDP. This function has been further analysed to a greater level of detail than the original CTA in order to provide specific, process-oriented MOPs that correspond to the major tasks performed by the Ops Room Team. This analysis was performed using a former Navy ORO who was part of the HSI team.

Note that no specific measures have been presently identified for functions relating new information to the Maritime Surface and Sub-surface pictures. We believe that the experienced gained in first measuring performance in the air domain and assessing the various possible measures will readily prepare us for subsequent T&E phases involving other warfare areas.

DETAILED MEASURES	COMMENTS
T&E TASK 1 Check mission picture 2.1.1	
1a Accuracy in detecting relevant info within incoming message stream	Manipulate info content relevant for circumstances, some immediate, some requiring later action. Vary workload Vary operational circumstances to influence salience of info Vary message source, TG, Ops Room, ship, GCCS Need SME judgement for ground truth
1b Accuracy in ignoring irrelevant info	Need SME judgement for ground truth
1c Total time spent in comms dealing with incoming info	Gross measure
1d Time to detect high salience message	For text – what specific event would be the marker for the start time. Assume for voice message this is not a problem.
1e Accuracy in requesting additional info (i.e. number of messages that required follow up that resulted in ORO request for more info)	Ensure that some messages require a follow-up enquiry. May be difficult to assess unless there is a large message pool.
1f Accuracy in ignoring lower priority messages	Ensure that messages vary in priority. May be difficult to assess unless there is a large message pool. Need SME judgement for ground truth
T&E TASK 2. Relate new info Ops Room picture 2.1.2	
2a Accuracy in directing (communicates) info/direction resulting from message	Could be voice or text action?
2b Accuracy re message content in briefing appropriate station	
2c Accuracy in comprehension of impact on pre-plans/response options/tactical situation	Need SME judgement for ground truth
2d Accuracy in recognition of impact on Ops Room capability	Need SME judgement for ground truth
2e Total time spent in relaying info re Ops Room status	
T&E TASK 3 Relate (new info) to RMP Includes aspects of build awareness across domains (RAP, MSP, MsubP). Same measures will apply in principle to *6. Relate to WAP*	
3a Accuracy of salient info within RMP *prior* to new message	Need ongoing regular probes of OROs knowledge of salient aspects of MTP.
3b Accuracy of salient info within RMP *after* new message	
3c Time for ORO to complete understanding of new info regarding RMP	T&E probe could be in the form of a CO standing order for a SITREP or pressing a function key when he believes new info changes tactical pic.

DETAILED MEASURES	COMMENTS
3d Accuracy in taking appropriate action as a result of updated RMP	Could provide info that requires specific action to be taken (e.g. change course, change direction to teams, advise CO/TG, review pre-plans etc. Hence we need to specify a wide range of salient information dimensions that will be manipulated.
3e Total time spent on SSD relating new info to RMP	Need to define a trigger event and "end" event - could require a specific T&E probe etc.
3f Accuracy in assessing the integrated tactical picture	Freeze probes or SITREPS. Includes knowledge of threats and contacts of interest in relation to ship or TG.
T&E TASK 4 Relate new info to RAP 2.1.5. (as an example): Ensure contact classified 2.4.3.1 Note: Same approach/measures to be used in principle for MSP and MsubP	
SUB TASK 4.1 Identify friendly aircraft Friendly / neutral a/c will provide the majority of items in the air contact stream. There will be two conditions: when no hostile a/c is being dealt with, and when one is being dealt with. Distraction / attention management factor will be quite different i.e. more likely to miss / short cut identification procedures when focussing on a hostile.	
4.1a Accuracy in identifying friendly aircraft (by type/mission)	
4.1b Mean time spent in locating friendly aircraft	Must be expressed in relation to aircraft load factor
4.1c Total number of queries or total time spent in querying team for additional info	
SUB TASK 4.2 Identify hostile/suspect aircraft Hostile / suspect a/c will be occasional insertions in main air contact data stream.	
4.2a Accuracy in identifying hostile / suspect aircraft (by aircraft type)	
4.2b Mean time spent in locating hostile/suspect aircraft	
4.2c Total number of queries or total time spent in querying team for additional info	
SUB TASK 4.3 Identify neutral aircraft	
4.3a Accuracy in identifying neutral aircraft	
4.3b Total time spent in locating neutral aircraft	
4.3c Total number of queries or total time spent in querying team for additional info	
SUB TASK 4.4 Identify NU tracks (non-updated)	
4.4a Accuracy in identifying NU tracks	
4.4b Total time spent in locating NU tracks	
4.4c Total number of queries or total time spent in querying team for additional info	
SUB TASK 4.5 Identify tracks reported by ownship or, conversely, by other participating units (PUs)	
4.5a Accuracy identifying air tracks being reported by ownship	
4.5b Total time spent locating air tracks being reported by ownship	
4.5c Total number of queries or total time spent in querying team for additional info	

DETAILED MEASURES	COMMENTS
SUB TASK 4.6 **Identify own force engagement status** This measure will only be used if scenario requires engagement	
4.6a Accuracy identifying air tracks being engaged by ownship or ships in TG	Or other friendly a/c??
4.6b Total time spent in locating air tracks being engaged	
4.6c Total number of queries or total time spent in querying team for additional info	
SUB TASK 4.7 **React to threat track symbology LINKed by consort not on video**	
4.7a Time to recognise symbology not on video	
4.7b Time to order remedial action	
SUB TASK 4.8 **React to LINK not gridlocked**	
4.8a Time to recognise LINK not gridlocked	
4.8b Time to order remedial action	
T&E TASK 7 **Assess threats -generic**	
SUB TASK 7.1 **Detect changes in tactical situation**	
7.1a Accuracy in ID new threats	Probe data to be inserted by T&E team
7.1b Accuracy in ID changes in threat status	
7.1c Time to detect new threat	
7.1d Time to detect change in threat status	
7.1e Time to provide SITREP	Request for SITREP provided by T&E team Can be done as freeze probe or as an ongoing request for info.
7.1f. Accuracy in SITREP contents	Need SME judgement for ground truth
SUB TASK 7.2 **Determine threat priorities across domains** Will need to vary circumstances to ensure that some threat changes occur in a particular domain while the focus is in that domain, at other times when focus is in another domain	
7.2a Accuracy in ranking threat priorities across domains	
7.2b Time to assess threat priorities across domains	
SUB TASK 7.3 **Assess threats- (air warfare specific) Identify air track threat priority** Similar MOPs should apply in principle to other warfare areas.	
7.3a Accuracy in identifying threat priorities	Accuracy in terms of time to intercept or aircraft weapon lethality SME to provide ground truth
7.3b Total time to identify three highest priority threats	
7.3c Time to annotate CCS with CPA	Determine closest point of approach (CPA) for each threat
7.3d Accuracy in determining CPA	RMS lat/long measure
7.3e Accuracy in determining lethality	Determine 'lethality' for each threat. SME to provide ground truth
7.3f Time to determine lethality	May need specific ORO action to indicate task completed
7.3g Time to create a SITREP	Appreciation of overall air tactical situation

DETAILED MEASURES	COMMENTS
7.3h Accuracy and completeness of SITREP	
SUB TASK 7.4 Analyse history profile[7] of hostile aircraft	
7.4a Accuracy in analysing attack history of hostile aircraft	Number of fighter/bomber (FBA) runs, or missiles fired
7.4b Time to analyse attack history of hostile aircraft	
7.4c Accuracy in assessing number of weapons remaining on hostile aircraft	Assume a given quantity at game start as assessed by intelligence sources
7.4d Total number of queries or total time spent in querying team for additional info	
7.4e Accuracy in assessing attack tactics used by hostile aircraft (individual contacts)	Firing range, altitude profiles, number of weapons used
T&E TASK 9 Assess sensors Sensor information can become degraded through equipment malfunctions or restrictions. Hence, will need to ensure that sensor becomes degraded during scenario	
9a Accuracy assessing that sensors are less than optimum	
9b Time to appreciate that sensor is performing less than optimum	Need to log time at when sensor becomes degraded and ORO takes action by looking for relevant ORO comm.
9c Accuracy in identifying current sensor range predictions	
9d Total time in identifying current sensor range predictions	
T&E TASK 10 Manage ship surveillance	
10a Accuracy in recognising the problems in detect to resolve process	Need to ensure errors of different types presented (e.g. amplification, inappropriate work focus, communication errors)
10b Time to recognise problems in detect-to-resolve process	
10c Appropriateness of remedial action to correct errors	SME evaluation

Table 3.4: Details of MOPs and comments on their implementation

3.4 Test and Evaluation Plan: Overall Strategy

3.4.1 General Strategy

A number of factors influence the general strategy for developing the test plan. These include:

- An initial focus of the COMDAT1 TDP on AWW
- Current uncertainty concerning the way the TDP will evolve to impact upon WAP information integration, and surface/sub-surface integration
- The need to acquire experience in developing scenarios for the specific test environments contemplated (NCOT, ORTT) and running T&E scenarios in the specific test environments. (Especially for NCOT which has not been used for such complex scenarios previously).

[7] By history profile we mean the patterns of trajectory shown over time by a particular contact that is potentially hostile; these patterns include changes in altitude, speed and direction, communication trends, EW emissions etc.

- Uncertainties over the specific logistical requirements for human-centered[8] testing, until pilot trials have been conducted

All of which suggest that an incremental approach be adopted for testing both in terms of logistical scope and complexity of the domain to be studied, in order to minimise and manage risk. We propose that the initial area of focus start with the RAP, in later phases we will look at the MSP, MSubP and the RMP. In each case we commence with simple scenarios to allow proof of concept methodologies and gain experience, and then progress to more complex scenarios in terms of events and players, to full Ops Room team simulations.

3.4.2 Choice of Approaches

There are two basic goals for all human-centred testing: first to collect robust and reliable baseline data for core and critical tasks, and, second to collect data for those same tasks as influenced by the TDP. There appear to be two possible approaches to data collection for the above purpose. First, to set up a T&E environment using one of the Navy simulation facilities and to conduct a series of trials specifically for the purpose of collecting MOP data. This report focuses on this approach and outlines the necessary detail to conduct the trials. A second approach would be to use existing data created by the Navy during training or work-ups. Our review of potential test environments showed that the ORTT facility is capable of running full operational scenarios involving the whole Ops Room team and recording in a high level of detail the actions, screen contents, communications of the team under study. We believe that the *records of such training sessions may prove to be a useful source of baseline information*, from which MOPs for a wide range of tasks could be extracted after the fact.

For example, by watching the RT1 screen during playback and following the AAW team communications it would be possible to time the detect to resolve cycle for a number of contacts of interest. Further, we would be able to monitor the ORO's interactions and interventions to ensure that processes are being conducted appropriately. We could also trace the ORO's actions in switching between different representations of the tactical environment on the CCS. Some of the measures extracted could be timed-based, others might be SME ratings or assessments of the ORO's actions and communications. Gross measures of communication could also be readily obtained as well as communication patterns between team members associated with specific scenario events. Our initial evaluation of the ORTT suggested that this information can be obtained with a sufficient degree of precision and reliability to meet the needs of T&E.

It will be recalled that the ORTT contains a group debrief facility with 3 large projections screens and loudspeakers for audio playback. In addition, data recorded during a training session may be replayed in the Training Control Facility. The system provides a debrief mode which allows replay of a simulation at one of four (0.5, 1, 2 or 5) play speeds for the debrief session of data that have been recorded. A Debrief Control is a tool available to the training staff to debrief trainees on completion of training sessions. It uses game data saved by the Game Monitoring functions. Three monitors, an audio system and a screen projector are available to debrief the trainees. Information including hardware panels, CCS data, synthetic environment, Link-11, CCS tactical pictures, trainee consoles and audio recordings can be presented. Editing tools are available to assemble and compile an audio/video presentation from the recorded data. There is a capability to playback the data at slower or faster than normal speeds and to jump to specific points of interest that have been tagged either prior to, or during, the scenario execution.

[8] The COMDAT1 SOR refers to this as human-in-the-loop testing (HIL)

Given the potential usefulness of the ORTT training records, *we recommend that a more in-depth evaluation be conducted to determine if useful MOP data could be extracted.* To achieve this, HSI® staff would conduct a site visit to the ORTT and review as a first step the types of scenarios and associated records that are available. The second step would be to select an appropriate sample playback and conduct a proof of concept analysis in terms of extracting potential MOP data. Access to such scenario records could be scheduled at off-peak hours[9] or during the eight hours per day set aside for software development. This site visit could be integrated with a review of an actual exercise in progress in the ORTT, the justification for which is outlined in section 3.6.9.4.

Since the potential usage of the ORTT to collect MOP data remains somewhat speculative at the present time, the balance of the section on T&E strategy will focus on providing details of a test program to collect real-time MOP data in a series of dedicated trials in NCOT.

3.4.3 Sequence of Major Test Trials

The above considerations suggest an approach that follows the following sequence of major trials:

1. Collect baseline performance on ORO tasks relating to the RAP. Measures 4,7.3,7.4,9, 10.
2. Repeat 1, as influenced by the TDP.
3. Collect baseline performance on ORO tasks relating to the generic detection of new information (i.e. not just air surveillance) and on ORO tasks relating to the MSP and MSubP. Measures 1,2,7,9 (adapted for other warfare areas).
4. Collect baseline performance on ORO tasks relating to the integrated RMP and WAP Measures 3. 7.1,7.2.
5. Repeat 3, as influenced by the TDP.
6. Repeat 4, as influenced by the TDP.

This sequence should be considered as having some flexibility depending upon the actual progress made in developing the TDP. For example, trials 3, 4 could be brought forward, if the technology were not in place to perform trial 2.

Each of the test trials will have a structure that involves at least the following elements:

- Logistical requirements
- Scenario development
- Detailed test plan
- Proof of concept
- Pilot study
- Main study
- Data analysis
- Conclusions and lessons learned

In the following sections, we outline some of these basic requirements for each of the above in as much detail as is currently feasible, given the current state of knowledge and experience.

[9] Note that the ORTT can only be used in either game playing mode or playback mode at any one time. Off peak access would therefore minimise conflict with Navy training requirements.

3.5 Logistical Requirements for Collecting MOPs

This section outlines the T&E infrastructure that will need to be in place in order to support the first test trial. One major uncertainty concerns whether all events can be pre-programmed into a scenario to run on a predetermined schedule, or whether live players of certain roles and "drivers" of data will be required. For example, in the case of ambiguous air tracks, it is not known in NCOT or the ORTT simulators, whether the kinds of operational behaviour of sensor systems found during actual sea conditions can be replicated. If not, live game players will be required to drive tracks in a particular way in real time in order to simulate real events[10]. Such uncertainties have major implications on the logistical overhead for the running of a scenario and the implementation of the scenario elements into the appropriate software. Such requirements will become better known after the first proof of concept trial.

In general, each trial will comprise following T&E phases.

3.5.1 Preparing the Trial

Naval liaison: This will be required to book the facility, arrange for T&E staff visits and access, arrange for test participants with appropriate naval SME background.

Scenario development and encoding: T&E personnel will require access to existing scenarios and related OPGENs and OPTASKs. T&E staff will need to be trained in scenario development and modification. If current OPTASKs are not available, Navy personnel will be required to assist in the development of this information. Software specialists will be needed to ensure that the scenario is pre-tested and functions appropriately. Access to suitable workstations may be required for T&E staff to proof and test scenario components. Naval SMEs will be required to provide some expertise on the events to be simulated (wherever possible this will come from personnel within the T&E team).

Software development and coding: Some modifications to the existing simulation software may be required to support T&E requirements. These might include the ability inject flags into the scenario database to signal start of key events, to provide a capability for injecting T&E information probes into the scenario in real time, to capture and log the time of participant actions at a workstation, to freeze the scenario and restart. The responsibility for defining the requirements lies with the T&E team, the responsibility for any software coding to incorporate the requirements will be with the software developers for a particular facility, for which suitable budgeting and contracting provisions will be needed.

3.5.2 Running the Trial

Information sources: these should represent all aspects of a normal Ops Room where the ORO can be expected to gain information. Sources of text (CCS or paper) and audio messages include the ongoing Ops Room team, the TG, other areas of the ship, the GCCS and stateboards and any other of the relevant communication nets.

Ops Room personnel-real or simulated: these represent "live" players with whom the ORO would normally communicate and interact. For most of the tasks anticipated to be impacted by the TDP, we would expect the major players to be the SWC, ASWC, CO and ORS. Information flow from

---

[10] As a result of a recent evaluation of the NCOT facility (see Annex B), we know that an RT1 will be required to enter radar tracks but that certain aspects of radar returns from contacts will need to be simulated using real-time control of game entities by the T&E team.

other members of the Ops Room including the front row team and sonar systems normally flows to the ORO through the warfare directors (WDs).[11] Based upon a recent, in-depth evaluation of the NCOT environment (Annex B), we believe that experienced Navy personnel will need to play the roles of the RT1 and SWC and/or ASWC when the trial involves other than air warfare. An experienced RT1 will be required to process the basic sensor data that comprise the radar picture that is used by the ORO and SWC.

In general, the dynamic information to be provided to the team that is beyond that contained within the scenario pre-scripted events and data will come from one of two sources, T&E personnel who follow a precise script and a Navy SME (part of the HSI® team). The latter will play many roles by providing all of the technical communication to the ORO (e.g. while acting as CO), by manipulating the tactical picture in real time as circumstances warrant, and by generally providing all specialised, knowledge-based information that cannot be pre-planned but is required for a realistic scenario. Another role for this individual will be to observe and make notes, for later analysis, on the actions of the ORO at certain times. The individual who plays this role must have an intimate knowledge of ORO functions, Ops Room procedures and expected performance for the particular scenario. For convenience, and given the omnipotent, all-knowing insight required of this role, we refer to them subsequently as the Gaming Operational Director (GOD).

Workstations: these are the physical simulations of actual Ops Room workstations that will need to be in place, in order to allow the associated Ops Room team member to perform their normal tasks. For the most part, we anticipate that most of the human-centred testing for the COMDAT1 TDP can be accomplished with workstations for the ORO, SWC and ASWC plus additional workstations for team members such as the RT1 or SCS, depending upon the domain focus of any trial. An additional workstation will be required for T&E personnel to monitor events on any of the subject workstations. This dedicated T&E workstation should have the capability to display the following information: CCS Data, CCS Tactical Picture, Link-11, Synthetic Environment from the perspective of any member of the operational team (usually the ORO). In addition, from 1-3 workstations will be required for GOD and other T&E personnel in order to be able to inject dynamic, real-time data and messages during the course of the scenario.

Environment: Necessary parameters of the operating environment will need to be simulated where appropriate. This includes selected elements from the natural, tactical, EW and acoustic environments. Normally this capability is inherent in the simulation facility to be used.

Data to drive the simulation: This will either come from pre-scripted events or will be provided in real-time by game players who act as data sources or drive game entities following pre-planned scripts for the most part. Some of the data will be provided by Navy SMEs playing the roles of the SWC, ASWC and subordinate members of these teams, as the scenario context demands.

Simulation support: this includes all personnel required to support the trial (other than role players) and will include network technical staff, software staff and naval liaison staff.

3.5.3 Analysing and Reporting the Trial

Data translation for analysis: If the output data from the trial is not available in a format that permits analysis in standard commercial database and spreadsheet applications, then support personnel will be required to provide this transformation capability.

---

[11] In later versions of COMDAT, when information related to process regulation and monitoring is developed, it may be necessary to have all Ops Room personnel represented. This is because a major role of the ORO in his management capacity is process refinement.

Data interpretation and analysis: This will be performed by the T&E team supported by Naval SMEs in situations where qualitative evaluations of the captured data must be performed. Facilities that may need to be provided for this activity include workstations for the replay of data and/or scenarios and equipment for the replay of audio or video logs.

Reporting the trial: The T&E team will have the responsibility for providing a written report to the scientific authority and to provide an oral briefing of the results. It is expected that this would be performed after completion of each of the data capture phases outlined in section 3.1 above.

The following table summarises the major logistical requirements for the four ORO functional areas to be evaluated.

	Detect incoming information	Build/Maintain RAP	Build/Maintain MSP,MSubP	Build/Maintain integrated RMP/WAP
Information Sources				
Text messages	X	X	X	X
Link 11 (CCS symbology)	X	X	X	X
External net TG	X	X	X	X
Internal net Ops Room: General (C&C)	X	X	X	X
Internal net team (AWW)	X	X	X	X
GCCS (symbology/map/text)	X		X	
Stateboards	?			
Ops Room Team Players				
SWC	X	X		X
ASWC	X		X	X
ORS	X			X
EWS				
CO	X	X	X	X
ORO	X	X	X	X
RT1	X	X		X
SCS	X		X	X
TS	X	X	X	X
RT2			X	X
Workstations Real/Sim				
SWC	Real	Real		Real
ASWC	Real		Real	Real
ORS	NA	NA	NA	
EWS	SIM	SIM	SIM	SIM
CO	SIM	SIM	SIM	SIM
ORO	REAL	REAL	REAL	REAL
RT1	Real	Real		SIM
SCS	SIM	SIM	SIM	SIM
RT2	Real		Real	SIM
TS	Real			SIM
SAC	SIM	SIM		SIM
HMS				SIM
Source of data for events				
Text messages	Pre-scripted/T&E team	Pre-scripted/T&E team	Pre-scripted/T&E team	TBD
Link 11	Pre-scripted/T&E team	Pre-scripted/T&E team	Pre-scripted/T&E team	TBD
Internal net Ops Room	Pre-scripted/T&E team	Pre-scripted/T&E team	Pre-scripted/T&E team	TBD
Internal net ship	Pre-scripted/T&E team	Pre-scripted/T&E team	Pre-scripted/T&E team	TBD
GCCS	Pre-scripted/T&E team	Pre-scripted/T&E team	Pre-scripted/T&E team	TBD
Stateboards	Pre-scripted/T&E team	Pre-scripted/T&E team	Pre-scripted/T&E team	TBD
Sensor data - air	Pre-scripted/live player	Pre-scripted/live player	Pre-scripted/live player	TBD
Sensor data - sub-surface	Pre-scripted	NA	TBD	

Table 3.5: Summary of logistical requirements

3.6 Detailed Approach to the Evaluation

An incremental approach has been suggested above for the conduct of test and evaluation and the collection of MOP data. This will ensure that risks are managed and resources applied wisely and minimally in the initial stages where feasibility is being confirmed. As more experience is gained in creating the necessary T&E environment and the collection of data, the scope of scenarios and range of MOPs may be expanded. The initial focus of the evaluation will be on measures relating to building and maintaining the RAP, since this is the function most immediately impacted by the COMDAT1 TDP.

Based upon our initial assessment of data collection environments, and a recent follow-up visit to NCOT to determine specific capabilities and limitations (see Annex B), *we recommend that the initial proof of concept and pilot data collection be conducted in the NCOT facility*. This environment appears to have the necessary capability for simulating air warfare involving the RT1, SWC and ORO as interacting players.

The first trial will involve the collection of MOP data for the function **"Build and maintain the RAP"** and will proceed in the following phases.

1. Generate software requirements for implementing proof of concept testing of environment and MOPs. Generate software.

This activity is expected to involve primarily HSI® staff and will result in the production of a requirements list to support T&E test trials for the software developers of NCOT. Any software that will be required to be developed to support the test trial (i.e. is outside of the current "build" capabilities of NCOT) will *need to be funded and contracted separately with McDonald Dettwiler.*

2. Generate scenario; encode scenario events in software.

By scenario we mean the sequence of events and associated contextual information that will form the basis for guiding the actions of Ops Room personnel.

HSI® staff will review existing scenarios that have already been created for training or other purposes to assess their suitability for T&E purposes with a view to modifying them as required. In the event that nothing suitable is available, then scenarios will be built from the ground up. Our preliminary assessment is that there is nothing currently available that has been pre-programmed for the NCOT environment to meet T&E needs, but a library of game entities is in place that can serve as the building blocks for generating scenario events.

In order to develop and test the scenario events, access to an NCOT workstation will be required for T&E staff. In order to reduce the travel overhead, inconvenience and other associated costs with doing this work at the NCOT facility*, we recommend that HSI® staff be given access to an NCOT workstation more locally*. This could be achieved by making an NCOT Unix workstation available at HSI® offices or at DRDC. Whatever the location, we strongly recommend that such a workstation be acquired, since the short term cost of acquisition will be more than offset by future costs associated with travel to Halifax. Further, the provision of such a workstation locally would better support the continuing needs of the development of MOPs related to COMDAT and would allow the local testing of concepts and scenario events in an efficient and cost effective manner. Another major advantage would be in providing a local capability to playback and analyse data from T&E trials, again without the overhead of travel to Halifax.

Notwithstanding the location of this workstation, the encoding of the scenario events into software and their initial testing will require the support of Macdonald Dettwiler staff, and contractual arrangements will need to be made to formalise and fund this process.

3. Initial proof-of-concept assessment and familiarisation with the selected test environment (NCOT in the first instance)

This process is designed to provide an early check on whether the scenario will run as required; whether the real time probes can be injected; whether events can be captured, timed and logged; to work out logistics for personnel who will be driving data and playing roles and to ensure that stored data are amenable to analysis. It will also be used to evaluate whether the appropriate level of realism can be achieved for simulated events such as ambiguous radar data and loss of radar data. Major tasks are the preparation of the scenario materials, on-site testing, analysis and reporting.

4. Revise scenario and data capture methods

Based upon the outcome of the proof of concept, some revisions may need to be made to the scenario elements and methods for generating and capturing T&E data.

5. Conduct initial pilot trial with selected MOPs/limited scenario

The goal here is to gather in a cost-effective manner some initial data from segments of the scenario to ensure that everything is working correctly before deploying the more extensive resources required for the main trial. All major forms of probes and data capture tools/methods will be represented. It is assumed that the pilot trial will comprise two days of data collection with morning and afternoon sessions. The personnel requirements for conducting the pilot trial are shown in the following table. As can be seen, three Navy SMEs will be required, of these the RT1 and SWC should be the same individuals for each of the four test trial runs. These role players will require to be additionally trained in the conduct of the trial, especially to act as T&E confederates under some circumstances (e.g. RT1 fails to amplify appropriately, SWC has wrong focus of attention). Further, training will ensure that having been through the same scenario trial on previous occasions that role players not give off inadvertent cues and maintain the same approach and procedures on repeated runs. A different ORO, who is the focus of the trial, will be required for each separate run.

The specific requirements for the pilot trial are outlined in the following table. Where a source is specified as SIM, this means that the information normally provided by that source will be simulated by making it available through T&E staff who follow a script, or in some cases by the GOD.

Information Sources	How Provided	Ops Room Team Players	Personnel	Workstations Real/Sim
Text messages	Pre-scripted/T&E team	SWC	Navy SME	Real
Link 11 (CCS symbology)	Pre-scripted/T&E team	ASWC	Not required	
External net TG	Pre-scripted/T&E team/GOD	ORS	T&E role player	SIM
Internal net Ops Room: General (C&C)	Pre-scripted/T&E team/GOD	EWS	T&E role player	SIM
Internal net team (AAW)	Pre-scripted/ Navy SME role players	CO	GOD	SIM
GCCS	Pre-scripted/T&E team	ORO	Navy SME	Real
Sensor data-air	Pre-scripted/software	RT1	Navy SME	Real
Tracks-surface/sub-surface	Pre-scripted/T&E team	SCS	T&E role player	Real
Stateboards	Pre-scripted/live player	TS	T&E role player	Real
	Pre-scripted	RT2	Not required	
CANEWS	Pre-scripted/T&E team		T&E role player	SIM

Table 3.6: Summary of logistical requirements for pilot trial

Analyse/report pilot trial

This will involve a full analysis of the data to determine the following:

- the scenario runs according to plan
- scenario events elicit the appropriate responses
- the responses are logged and recorded appropriately
- game players can fulfil the task roles
- the captured data can be analysed and provide the right kind of information for T&E purposes.
- Preliminary estimates of MOP data ranges and variability.

The trial will be reported to the Scientific Authority as a written technical report and an oral briefing.

6. Refine MOPs and scenario

On the basis of the lessons taught from the Pilot trial, modifications will then be made to the scenario and/or MOPs and possibly the T&E software requirements. Since it is unlikely that all potential MOPs can be assessed, given logistical constrains on the experimental design and availability of resources (see below), a selection of the most salient and meaningful MOPs to include in the main study will be made.

7. Conduct second pilot trial with more complex scenario, if required.

Depending on the degree of success with the first pilot trial, this step may or may not be necessary before committing the full resources to the main trial. For now it has been included in the schedule as a safeguard.

8. Analyse second pilot/Refinement of MOPs and scenario

Same as for items 5 and 6 above.

Note: steps 7 and 8 may not be required if the outcome of the first pilot trial is successful, or if any required modifications are minor in nature and will not require to be formally tested again in NCOT or the ORTT. These steps are seen as providing a conservative estimate of a worst case situation in order to plan for resource allocation and project timeline.

9. Build and proof scenarios for main trial

The complete scenario for the main trial is completed based upon the information learned to date. The full scenario is tested and rehearsed using T&E personnel to simulate Navy roles and to drive data as required.

10. Conduct main trial/ collect baseline data

This is the formal data collection trial. The scope of the trial is determined by a number of factors relating to data reliability as outlined in a subsequent section. It represents the appropriate level of effort to establish reliable baseline performance data for the major ORO functions.

The availability of the OROs is a major concern for being able to conduct this trial over the consecutive sequence of days proposed. We believe that a minimum of eight data sets be captured, each using a different ORO. Plus we should add an additional ORO for contingencies. Given the limited pool of OROs, obtaining such a large sample at any one time may prove to be difficult. An alternate approach that may be considered, if the availability issue cannot be solved, is to consider running four ORO's with each performing two separate sessions. Clearly, in this case different scenarios will need to be prepared for the two different runs, and this will require more preparation and development resources by the T&E team than has been determined in this initial estimate.

11. Analyse and report main trial

A significant level of effort will be required to thoroughly review all of the data captured. This will include not only events and responses captured by the software, but also any video or audio records and paper message traffic.

12. Refine measures and scenarios.

Prior to the conduct of Trial #2, further refinement will need to be made to the MOPs and scenario to accommodate emphases on different warfare areas. The overhead for preparing, proofing and piloting subsequent trials should be somewhat less than the first trial because of lessons learned and experience gained.

The next table provides an initial approximation of the human resource requirements to accomplish the above. Not included in this table is the proposal to visit the ORTT to review existing scenario records for potential MOP usage (see section 3.2) and to observe an exercise in progress (see 5.7.4). It is estimated that these two activities would require 10 HSI® person days.

The estimates provided below include time for travel to Halifax to conduct the relevant activities and assume that a workstation is not available locally for scenario development. The SWC and

RT1 Navy role players should ideally remain the same throughout the pilot trial and also the main trial, although not necessarily the same individuals on the two occasions. For the main trial we have built in an extra half day cushion to allow an additional session to be added in case of problems arising that result in the loss of a test session. A more comprehensive and detailed estimate of the hours, task allocations and costs to conduct all activities up to the pilot trial has been provided separately to the Scientific Authority.

Task	HSI® staff	Navy SME (normally from HSI® staff)	Simulation facility software developers	Simulation facility support staff	Navy SME role players	Navy SME test participants
1 Generate T&E software requirements. Code software	3	1	Unknown level of effort			
2 Generate and encode scenario	5.5	3				
3 Proof of concept						
3a Preparation	2	2		Unknown		
3b Conduct	4	1	2			
3c Analysis/reporting	3.5	.5				
4 Revise scenario/methods	3	2				
5a Conduct Pilot (2 days of testing)	9	2.5		Unknown	SWC-2 RT1-2	4 (ORO)
5b Analyse pilot	7	2				
5c Report pilot	8.5	1				
6 Refine MOPs/scenario	5	1				
7* Pilot 2 - 2 days of testing	5	3	1	Unknown	SWC-2 RT1-2	4 (ORO)
8* Pilot 2 Analysis	2	1				
9 Build and proof scenarios for main trial	2	2				
10 Conduct Main trial (assume four days with am/pm sessions plus half day set-up, half spare	20	5		Unknown	SWC-5 RT1-5	8 (ORO)
11 Analyse and report main trial	20	5				
12 Refine measures & scenarios	8	4				

Table 3.7: Approximate personnel resource requirements for each trial phase (numbers are estimates of person days.) *Note: these steps may not be required**.**

3.6.1 Issues Concerning Data Reliability

Given the level of effort required and significant resources deployed to collect baseline data, the T&E team need to make every effort to ensure that reliable data are generated to form a baseline for future comparative purposes. A number of sources can be identified that will affect the variance of collected data and thereby provide constraints on its reliability and generalisability. These constraints are listed below together with an assessment of how they will need to be treated in the T&E trials.

3.6.2 Subject Variability Among ORO's

There are a number of factors that will influence the performance of different OROs. These include range and depth of experience, individual abilities, individual motivation and recent familiarity in doing the ORO functions required in T&E. Variability in each of these domains can seriously widen the confidence limits around mean data. In order to address these issues a combination of selection constraints and choice of appropriate numbers of test subjects is required. Therefore, we suggest that selection be limited to currently active ORO's with a minimum of one year of operational experience.[12] Selection may also include OROs who have been out of operational service for less than one year but have three or more years of prior operational service. A minimum sample size of 8 ORO's will be required for each test trial in order to provide reliable estimates of inter-subject error variance and to obtain sufficient statistical power to detect any performance differences resulting from the TDP (see 5.2 below). Further, it is preferable to use a different sample of subjects for Trial#2, in order to avoid any potential carry-over effects from Trial#1.

A major practical issue that must be considered is that on the East and West Coast combined there is a theoretical maximum of about 45 OROs who are currently in active service, or have had a tour of duty at sea within the last 12 months. However, because of sub-optimum manning levels, this figure is more likely to be closer to 30. Hence, there may be stringent limitations on the availability of appropriate OROs to participate in T&E trials.

3.6.3 Variance Associated with ORO Workload Factors

OROs operate under a variety of levels of workload. Under high levels of load they frequently switch from task to task, spending less time on each than they would if they were underloaded. In order to assess the generality of any improvements in performance that may result from MSDF technology, it will be important to sample from work situations involving different load levels. Further, because OROs may be working at almost optimum performance when underloaded, there may be a ceiling effect on performance that reduces sensitivity in detecting any further improvement in performance due to MSDF. This suggests that devising a test scenario in which ORO's focus only on a subset of critical tasks without any significant workload loading will yield data of potentially minimal value. Thus, it will be necessary to ensure that tasks result in a sufficient level of workload that ORO's are not operating at their maximum capability.

Workload loading for the ORO is not necessarily a homogenous variable that can be characterised by a simple quantitative measure. Workload may vary in terms of the volume of data within a task domain (i.e. ranging from a small number to a large number of air contacts). It may also vary in

[12] Of course, if the Navy is interested in collecting data on how experience affects performance on these tasks, it will be necessary to select two groups of participants who differ in mean years of experience.

terms of the complexity of the data within a domain, for example many easy to identify air contacts that are widely dispersed represents a much lighter load than that same number of contacts who are unknown, or enemy and clustered along vectors and altitudes. Further, the number of ongoing tasks across domains will also influence workload, for example when the ORO must co-ordinate responses to potential air, surface and sub-surface threats. The sources of workload variability that may influence ORO performance are outlined in the following table. For within-domain sources, only information relating to the RAP is elaborated, since this is the area of initial priority, and the other domains will follow a similar pattern.

SOURCE OF WORKLOAD	WORKLOAD FACTORS
Within domain	Number of simultaneous contacts Rate of contact Number of unknowns ID ambiguity Track ambiguity Path predictability Number of lost tracks
Between domains	Number of domains (air, surface, subsurface) Differences in volume of data for different domains Number of threats across domains Priorities of threats across domains
Process monitoring	Number of ongoing processes Personnel experience and resource levels Equipment problems Demands from TG Demands from CO

Table 3.8: Sources of ORO workload

Given that MSDF technology to support process monitoring does not fall within the scope of the COMDAT1 TDP, the focus on workload manipulation for present purposes should be on the other two sources. In order to provide a representative range of workload, it is proposed to have three levels: two levels of workload would be achieved by varying within domain factors and the third by introducing workload from another domain. While it might seem initially appropriate to concentrate on manipulating load factors solely within the air domain (since the TDP will largely centre on air), we believe that this would be an error. Under conditions of high workload in the air domain and in the absence of any other task demands, the ORO would focus exclusively on air warfare. However, this would not be representative of actual operations, where the ORO has always other tasks to timeshare, and would therefore result in some overestimate of ORO performance. However, if we supplement a high within-domain load by a moderate load in another domain, or from other concurrent tasks, the resulting loading will have greater external validity to the operational situation.

We propose that within the overall trial design that **three levels of workload** be sampled as follows:

Moderate workload: The ORO and team are working at a steady but comfortable pace that can be easily sustained, and can cope with the rate of information. The majority of contacts come from the domain of interest, however there is a constant but low level of contact information from other domains.

High workload: The ORO and team are reaching the point of overload; they are working at a high pace that causes some stress if sustained; they can barely cope with the rate of information and some tasks are truncated or dropped; errors may be made. This high information rate is confined to the domain of interest, the contact information from other domains would be sustained at the same level as in the moderate workload condition.

High workload+ extra-domain loading: As above, plus the type and volume of information in the other domains reduces the ORO's capacity to focus largely the air domain. The additional information may involve dealing with potential surface or sub-surface threats.

3.6.4 Research Design Trade-offs and Sample Size Considerations

One of the major concerns in conducting this form of T&E activity is the constraint imposed by the availability of the facility, support personnel and participants. Unlike an environment specifically designed for research, with a dedicated complement of support personnel and readily available subjects, the primary purpose of NCOT and the ORTT is for Navy training. It seems likely that T&E opportunities will be limited in frequency and duration. As a consequence, the research design must be tightly focussed and sampling procedures must be highly efficient. Trade-offs will have to be made about the extent to which a comprehensive data set can be gathered for establishing a database of baseline ORO performance and to evaluating the effects of the TDP. One way to approach this problem is to consider the number of individual data points that will need to be captured and then work backwards up through the design, to see what is feasible in the likely time to be allotted and personnel (test subjects) to be made available.

The best way to approach this problem is to address issues of the expected magnitude of effects of interest, the anticipated error variance, the acceptable probability of making a Type II error, and the statistical power required. By providing *a priori* ranges of values for these parameters we can readily determine the number of data trials that will be required.

Another factor to be considered in estimating the number of data points required is whether any given MOP will be time based or accuracy based. In practice, accuracy based MOPs require far more trials to achieve the same size of equivalent error variance than response time MOPs. Ten measures of response time per individual will give a reasonable estimate of mean and variance. Whereas ten measures using proportion of items correct can easily be influenced by outlier performance on one or two trials. Further, one would expect that for most ORO tasks relating to situation awareness and communication, training ensures that performance operates at a high level of accuracy. If this is the case, detecting any differences in performance attributable to the MSDF TDP will be difficult to impossible, because of the existing performance ceiling. Therefore, the general focus on evaluation will be to use time-based measures, supplemented by accuracy measures, whenever there are clear instances of tasks that produce consistent errors in performance.[13]

3.6.4.1 Magnitude of Effects

We have no advance indication of what expectations the Navy may have concerning the effectiveness of MSDF in improving performance. Is a 5% gain of operational significant?-possibly not, unless the task is repeated with high frequency. Performance gains of 10% or more

[13] This general argument does not apply to those MOPs that will require subjective analysis of ORO actions and responses by SMEs, typically for ORO functions involving situation analysis and decision making.

are likely to result in increased efficiency that translates into increasing the ORO's spare capacity. Since we cannot anticipate Navy expectations in this regard, *we recommend that the Scientific Authority in discussions with the Navy should address this issue*. Additionally, we could have SMEs review the display concept during prototype development and estimate the magnitude of the changes in performance that might be expected for different tasks. The data generated by this process could then be used to focus the MOP effort. For example, in areas where large changes are expected, or where the changes are too small to be of operational interest, then there would be no need for a large concentrated T&E investment, instead the T&E effort could be better directed to areas where the impact was less certain.

In the interim, we will proceed with determining estimates of required sample size by exploring the implications of effect sizes of 10% and 20%.

3.6.4.2 Expected Error Variance

Normally estimates of error variance (that are used to guide sampling decisions) are obtained from the relevant literature or pilot studies. In the present circumstances neither of these sources appear to provide any useful guidance. We do not know how long some of the ORO tasks might take, or the variability underlying them. Therefore it seems appropriate to generate a possibility matrix that covers the range of circumstances that may be anticipated to guide the decisions on required sample size.

Based upon our knowledge of the ORO functions, observations of exercises and familiarity with C2 operations in other domains, we propose that the lower range of task completion times of interest is probably 5 seconds. This may seem too short a time to consider from the perspective of achieving meaningful operational increments through MSDF. However, if the tasks that produce this kind of response latency are highly frequent occurrences, small saving in efficiency will accumulate to the point of being operationally significant. At the other end of the range, we can only make a guess as to how long some tasks may take. Arbitrarily we must make some cut off, otherwise if an event produces a typically response that may take 20 minutes to evolve, there will be insufficient time in any one test trial to have adequate repetitions of such events. For working purposes, we have chosen an upper limit of 5 minutes as being the longest response latency that we can effectively deal with. This then gives us a range of potential mean response times between 5 seconds and 5 minutes. We can then interpolate some values between these limits and look at the impact of the range of expected means on sampling requirement by taking into account anticipated variance around these means.

For the purposes of generating some idea of sampling needs, we recommend as a starting point considering standard deviation values that represents 10, 20 or 30% of the mean. These numbers are based upon what might be typically found in complex reaction time or search experiments, although sometimes standard deviations that are 50% of the mean are found. However, such large values impact severely on the ability of the study to be sensitive to differences of interest.

The magnitude of the effect of interest and the error variance can be used to calculate a standardised effects size (**d**) which is defined as (Mean 1 - Mean 2)/ SD. Using this definition, in the standard psychology literature, **d** values of 0.2, 0.5 and 0.8 are regarded as small, medium and large effects, respectively.

Taking our working example of a 5 second RT and differences of interest of 10 or 20%, with estimates of the SD as being 10, 20 or 30% of the mean value, we arrive at the following table of **d** values. As can be seen, these estimates represent extremely large values for d, well beyond what is

regarded in the literature as being large. However, we can use the desired magnitude of the difference and the estimates of the variance to estimate the sample size requirements as outlined below, once we have discussed issues of statistical significance and power.

Possible mean value (seconds)	**5**	
Magnitude of effects	**10%**	**20%**
SD as a proportion of the mean		
10%	1	2
20%	0.5	1
30%	0.33	0.67

Table 3.9: Values of d computed for various assumptions concerning effects sizes of influence for test case RT = 5 sec

3.6.4.3 Desired Level of Statistical Significance

Typically, research in the social sciences sets the upper limit for accepting that the observed results are due to chance (%)= .05%. This means that five times in a hundred we will incorrectly conclude that when there is no effect, we will say there is one. This level may be too stringent for the kind of exploratory research and investigation that will form the TDP evaluation trial, and to establish a set of baseline performance measures. Here the goals will be to overcome the many sources of error variance that could work to make the study less sensitive to differences of interest, and maximise the chances of detecting any potential change in performance that could be of operational importance. Hence, we recommend a slightly less stringent level for making Type 1 errors by setting %=.1. In practice, this will mean fewer trials will be required to achieve a statistically significant outcome.

3.6.4.4 Desired Level of Power

Statistical power refers to the ability of the design to establish performance differences of interest, or the odds of confirming a theory correctly. Obviously, one would want this to be as a high as possible within the realms of what is achievable and practical within the constraints of time and effort. Too little power will result in a situation where the study has little chance of detecting significant effects. Too much power will mean that too much data are generated to the point that trivially small effect sizes are detected. In recent years a common, yet arbitrary, choice for a power level is .8

3.6.4.5 Implications of the Above for Sample Size

Having selected approximate values of the magnitude of the effects we are interested in, the range of anticipated possible mean values (for response time MOPs), the range of error variance and established significance levels and the desired power, we are now in a position to determine the impact of these variables on the required sample size. The following table shows the required sample sizes for the range of values outlined above. The top row provides three different possible values for means, the second row shows the size of difference between means that would be of interest and the body of the table shows the sample sizes that would be required under three different assumptions about the size of the standard deviation. A further assumption is that differences will be assessed using a two-group F-test (analysis of variance).

Possible mean value (seconds)	5		20		40	
Magnitude of effects	10%	20%	10%	20%	10%	20%
SD as a proportion of the mean						
10%	8	6	8	6	8	6
20%	28	8	28	8	28	8
30%	58	16	58	16	58	16

Table 3.10: Estimates of required sample size to reach required conclusions

Note that whatever we select for the estimate of how long the RT may be, the estimates of sample size all come out to be the same.

To interpret this table let us take a mean baseline response time of 20 seconds, and assume that a difference due to the TDP of 10% would be of interest, this would require a sample size as low as 28 if the SD were about 4 (i.e. 20% of the mean), but a sample size of 58 if the SD were about 6.7 seconds (30% of the mean). The detection of a 20% difference between baseline and TDP would require samples of 8 and 16, respectively for the same variance assumptions. Two, somewhat obvious, but important general principles follow from this, and should be always remembered in considering the design of the study. First, smaller effects are harder to detect and require larger sample sizes. Second, higher error variance reduces the ability to detect effects of interest and also increases sample size needs.

Based upon the above, a reasonable goal for the study would be to try to detect differences of 20% and to keep standard deviations around 20% of the mean value. If this can be achieved then we will need a sample size of 8 participants for each condition (i.e. baseline and baseline plus TDP).

For the present, we should bear the above number in mind in considering the overall demands on the time of ORO participants. Clearly we are faced with a number of unknowns. First, we do not know the extent of the ranges of the independent variables or how many levels of each will be required. Second, we have no estimates of the kinds of performance levels we can expect of the ORO participants. Third, the extent of inter-ORO variability on task performance is unknown. Fourth, we do not know the capability of the system to collect reliable MOPs data. *Consequently, it would seem prudent to not go into detailed design issues beyond Trial 1 at present*. Once this trial, or even the pre-cursor pilot trials to these have been completed, we will be in a safer position to outline the specific requirements for the subsequent trials.

3.6.5 Requirements for Scenario

The following table outlines the major elements that will comprise the scenario.

Geographical context

- Littoral environment between two large land masses that comprise Nations A and B, approximately within 30 nm each side of ship
- Land masses have mountainous areas coming close to coast that create hills and valleys[14]

Political context

- Nation on left and right land masses are potentially hostile to Canada/Allies and hostile to each other
- Tensions are high between both countries
- There are threats by both sides to embargo international waters to oil exports

Military context

- Halifax class ship operates in a Navy TG comprising: a high value unit, one Iroqois class and two other Halifax class frigates.
- A US battlegroup is within 150 nm of the Canadian TG. Intensive carrier based flight schedules are ongoing. These may generate between 10-20 friendlies in the sky at any time (10 for baseline workload; 20 for high workload)
- Nations on left and right each have airfields within 10 nm of coastline
- Nations on left and right conduct regular air training ops that include simulated air combat and bombing runs

Background theatre

- Area is in vicinity of two major commercial air lanes. One is medium-high-level with a/c largely in transit, the other is for a/c that are landing and taking off from an airport located close to one of the military bases.
- Local helicopter traffic to and from oil rigs.
- Local surface traffic comprising a mix of small vessels (e.g. media, recreational), commercial traffic (e.g. fishing vessels, tankers, and bulk cargo carriers), fast patrol vessels belonging to adjacent nations.

Op Orders

- Range of tactical area of interest (AOI) is a radius of 125 nm
- No pursuit or engagement of threats

General level of commercial air traffic

- For baseline workload- 30 commercial a/c in AOI - appearing and dropping off at a rate of about 1 every 60 sec
- For high workload - 45 commercial a/c in AOI – appearing and dropping off at a rate of about 1 every 45 sec

General levels of threat/unknown

- For baseline workload- 2-5 fighter bombers at any one time
- For high workload - 7-10 fighter bombers at any one time

Table 3.11: Scenario Requirements

3.6.5.1 Approximation of Number of Events per Test Session

In order to make the scenario and associated tasks in building the RAP realistic, we should conform to the kinds of rates of information that might be expected in actual operations as well as an appropriate proportional mix between non-threat and threat events. Thus, we would not expect that all contacts presented during the trial to be contacts of interest, nor will every contact be enhanced by MSDF. It would be unrealistic to have all contacts as unknown/possibly hostile, since this would bias the performance of the ORO and air team in a way that does not normally occur in either real operations, or in training simulations. Instead, the targets of interest (from the point of view of the evaluation of the TDP) have to be embedded within the normal stream of contacts that would be encountered. The question then is what would be an appropriate rate. The practicalities of the design requires a high number of contacts (since we do not care much about the contacts of non-interest), yet realism and the avoidance of bias demands a more moderate rate. As indicated

[14] This particular environment is known to produce track identity problems for radar systems that are primarily designed for open water. The map database for NCOT is not able to generate this degree of land mass complexity, hence the behaviour of radar under these circumstances will be simulated using real-time control by the T&E team of the game entities underlying the radar tracks.

above, we suggest that non-threat contacts occur at a rate of one per 60 seconds up to a maximum of 30 on the screen at any one time for the moderate workload condition. They also drop off the screen at the same rate. For the high workload conditions contacts would occur every 45 seconds up to a maximum of 45 on the screen. For a two-hour trial, there would be a total of 120 and 160 contacts for the moderate and high workload levels respectively. We suggest that we superimpose on this a rate of between 25-35% for the contacts of interest as a percentage of the overall number of contacts. This will mean about 30-40 (moderate workload) or about 50-60 (high workload) trials that will be available for MOP data collection. Given that a sample size of about 10 repeated trials for each measure will be required, this suggests that the design will handle between about 4-6 different types of events involving different behaviours by contacts of interest.

In addition to accommodating the main session in which data will be collected, the design will also need to provide an initial period of training for the participants. It is suggested that 30-45 minutes at the start of the test session be allocated to this.

3.6.5.2 General Structure of Test Trial and Course of Events

1. **Pre-watch scenario build-up**. Prior to the start of data collection, the background picture is built up by the T&E team. This will probably take 10-15 minutes.
2. **Watch handover:** ORO is briefed - about 5 minutes
3. **Phase in**: Routine watch activities to allow a period of relative peace and quiet for ORO to become aware, settle, and for T&E team to collect routine new contact cycle data. Air situation =25 commercial, 4 friendly in AOI (Note in high workload/across domain condition will also need starting scenario for MSP, MSubP)- about 10 mins
4. **Main scenario events:** this will take about 90 minutes and have the following features.
    - Commercial traffic added and dropped continuously
    - Adjacent nations A and B are conducting ongoing air ops with take-offs every 5-10 mins, circuits and occasional flights towards each other.
    - Unknown/possible threat air ops involve a/c doing some (not all) of the following event types[15]
        - simulated bombing runs on range
        - flying through hills and valleys
        - flying in circuits that progressively get closer to the TG
        - close formation and changing formation
        - attempting radar stealth
        - two a/c on same course/speed in close formation- one at constant altitude the other descending/ascending
        - two a/c that converge/run together for a time, then diverge
        - two a/c in close formation on a general heading to overfly the ship; a/c criss-cross with increasing frequency as they get closer
        - three a/c in close formation with one breaking off and criss-crossing the path of the remaining two

[15] These are based in part upon the a/c flying patterns used for the ASCACT trials and are believed to create problems for current sensor systems

- missile separation from one a/c (i.e. when a new air target appears suddenly and close, a pop-up target)

Given limitations outlined in the previous paragraph in terms of number of trials/events of interest, some prioritisation of these event types will need to be made.

3.6.6 Methods for Data Capture

3.6.6.1 Software Capture

The software will be required to capture and record the time (preferably to nearest .5 sec) of specific trigger events during the course of the scenario[16], the full details of these will be worked out during the detailed scenario planning and initial proof of concept trials. The events to be captured will include: specific key presses by the ORO and other Ops Room team members (where applicable) and selected key presses from the T&E team console(s).

3.6.6.2 Video and Audio Capture/Analysis

The system will be required to capture and record the time (to nearest 1/10 sec) of all audio and text based comms including their content and the recipient/sender. A complete audio record will be maintained on tape (or equivalent) of all comms, including those involving the Ops Room team and those between the Ops Room team and the T&E team. ORO comms that are direct and do not go through a network will also need to be captured. This may require a separate microphone on the ORO attached to an audio tape recorder.

The system will be required to capture screen contents from the ORO CCS on a time base accurate to 1/10 sec. The specific requirements will be worked out during the detailed scenario planning and initial proof of concept trials. The resolution of the screen capture shall be sufficient to be able to discriminate track details, symbology and all text.

3.6.6.3 T&E Probes

The system will allow the T&E team a capability to send from a remote terminal a message to the ORO in one of three formats: audio, screen based text, or paper-based text (hand delivered). A time stamp at the moment of issuance of the audio and screen based messages will need to be recorded. Any response to the message required by the ORO that will use the CCS or audio system will also be time stamped.

The system will allow the scenario to be stopped (and the tactical data on the ORO screen may be required to temporarily eliminated) to allow the T&E team to conduct probes of the OROs knowledge of screen content. These probes will either be by audio communication or in the form of a message sent to the ORO console from the T&E console. The system will allow the scenario to restart with full data intact with no scenario elapsed time at the conclusion of the T&E probe (likely within 3-4 minutes maximum).

[16] If the existing software cannot be modified to perform such timing, then the intervals of interest will have to be determined after the event by having the T&E team scrutinise and analyse the playback record. This will create a considerable additional burden of analysis that has not been factored into the interim estimates of personnel resource requirements.

3.6.6.4 SITREPS

To meet some T&E purposes, the ORO may be required to provide a brief SITREP**.** Options for the SITREP include either a verbal report (recorded on audio tape or by the software) or a written "fill in the blanks" report. We favour the former as it is more natural, less time consuming and does not provide the kind of contextual prompts available in the written format. The software will need to record the time of the request and time of completion of a SITREP and the screen contents of the ORO workstation at the time of the request (or any other screens that the ORO uses in order to comply with the request).

3.6.6.5 SME Real Time Observation

In order to evaluate some of the more complex ORO behaviours, we recommend the use of an experienced ORO stationed in close proximity to allow observation of the ORO test participant, much in the same way as training is now conducted. The role of the SME will be to evaluate ORO performance along prescribed criteria to be developed during the scenario construction phase. The observer will be provided with a workstation that can display either the same content as that of the ORO or any other information currently capable of being generated by the scenario. The workstation will also have the capability to capture selected key presses of the SME and allow messages to be created and notes to be recorded.

3.6.7 Data collection and Management Tools

It follows from the above that the following types of data are to be collected and that each will have its own requirements for management. In general, it would be desirable if all forms of critical data for analysis are provided in a medium or format that allows them to be taken off-site for analysis on a standard Microsoft Windows based platform (assuming that there are no security issues).

3.6.7.1 Response Time Data

These data are collected by the simulation software using prescribed trigger events or flags initiated by the T&E team. An underlying timebase with a resolution of 0.1 sec from the start of the trial must be maintained in order to mark events. Any individual response time is terminated by an event determined during scenario construction.

Response times (RTs) may also be required for events that will require a verbal response by the ORO, in these cases, the logging of such responses will also require a running timebase. (See also audio communication below).

The software will allow RTs associated with events to be coded in such a manner that they can be quickly retrieved and organised after a test session is complete. It would be preferable if this retrieval process could be expedited on the same day as the trial and be handled by the T&E team, rather than requiring the data to be sent to the software contractor for extraction. The data extracted by this process will be imported into an Excel spreadsheet where it may be organised appropriately for analysis.

3.6.7.2 Accuracy Data

Accuracy data may relate to several different kinds of information. These include the accuracy in knowing and responding to screen content (e.g. location and meaning of symbology, tactical data of importance), accuracy in making tactical estimates (e.g. closest point of approach), accuracy in communicating, accuracy in situation assessment and decision making. Given the variety of the

information content in all of these possibilities and the different modes of response, a single prescription for the collection and management of accuracy data cannot be provided. In some cases the information will be retrievable from ORO key presses or SSD screen-dumps, in other cases from text or oral communications, in other cases from responses to T&E probes (either in real time or if the scenario is temporarily frozen), in other cases it may have to be provided by the observer SME in real time or post event scenario playback.

The process of determining the accuracy of responses will be largely a manual, post-event analysis conducted by the T&E team on the basis of replaying the appropriate segments of the scenario, examining the events of interest and manually recording the responses that were made. In other cases, accuracy data will be provided from the SME observer in either interval data form or as a behavioural rating on a pre-determined scale. Whatever method is adopted, the resulting data will also be entered into an Excel spreadsheet.

It follows from the above, that the ability to replay, fast forward, slow down and halt a scenario record is a pre-requisite for this type of analysis. This requirement applies not only to data captured by the software but also to audio and video data.

3.6.7.3 Audio and Video Data

Two forms of video data will be captured - screen contents on workstations of interest and a video record of selected players. The screen content data will be captured by the system software and should be amenable to post-event analysis by replay on a workstation. The replay should allow a running timebase to be displayed at all times. An ability to extract single images of screen dumps during replay would be desirable. These should be in a graphics file format that allows them to be viewed with standard commercial graphics software. A JPEG format would be preferred to maintain file sizes at a manageable level. These data should be capable of being exported to other systems outside the T&E environment.

The video record of selected team players will be capable of being played back on standard commercially available equipment and will provide an on-screen running timebase.

The audio record will comprise all net-based communications and those captured directly from the ORO's microphone, that do not go through the standard comm channels. The audio record should be capable of being married to the video record where appropriate.

Audio and video records will be managed by maintaining a log and database. Copies of all critical records will be made for back-up purposes.

3.6.7.4 Observer SME Ratings

To review, SME ratings may be recorded in real time during the course of a trial or may be generated during post-trial scenario replay. In both cases, the data will be entered into an Excel spreadsheet in a format that allows events and their associated evaluations to be correlated. All subsequent analysis of the ratings will be maintained within the spreadsheet and associated with other MOP data where appropriate.

3.6.8 Data Analysis

The tools for data analysis will either be the resident statistical analysis functions in Excel, or where these are insufficient, the data will be exported to SPSS for analysis.

Data for each measure will be collapsed over samples and a representative statistic applied for central tendency. This will normally be the mean, but could be the median for data samples that are highly skewed (e.g. response times). Confidence limits around each mean will be calculated.

To compare baseline with baseline+TDP performance, an overall multivariate analysis of variance (MANOVA) will be used that includes all relevant MOPs. Subsequently, univariate ANOVA, t-tests and planned comparisons will be used to assess any differences in performance on individual MOPs. Magnitude of significant effects will be expressed as a percentage increase or decrease in performance. Should high data variance occur across subjects that masks the identification of significant effects, statistical tests might be performed either on difference scores (TDP-baseline), or by using a non-parametric sign test (e.g. TDP-baseline is expressed simply as a plus or minus depending upon the direction of the performance difference).

3.6.9 Constraints/Limitations/Risks

3.6.9.1 Operational Realism

The emphasis on designing the T&E trial has been to maximise the opportunities to collect robust and reliable data. As a result, some operational realism may be lost, for a number of reasons.

- The workload associated with the processing of unknown tracks and contacts of interest may be higher than that experienced operationally or in training. As such, when conducting the first trial (i.e. centred on the RAP) the ORO may focus more on air operations than would be normally the case.
- Many of the other factors that contribute to workload and the distractions that occur under normal operational circumstances will be absent.
- In real operations there may be considerable lulls in the level of background air traffic and in the rate of appearance of contacts of interest. Thus, workload may at times be somewhat low and then followed by a period of more intense activity. The constraints resulting from maximising the opportunities to collect data of primary interest means that such lulls cannot be afforded in the scenario event sequence.
- Although T&E participants will be thoroughly briefed and have a warm up period, by comparison with operational reality they will be dropped "cold" into a busy scenario with which they may have little familiarity. This is unlike the operational context where accumulated experience over previous watches will serve to guide and influence performance on the current watch.
- In reality, performance by the ORO or other relevant members of the Ops Room team is shaped by the cumulative experience of working together as a team. However, in the T&E trials such team cohesion will be absent, as the ORO will largely be working with individuals for the first time.

The first two of these factors suggest that, the T&E trial may overestimate ORO performance levels compared with what might be expected in operational conditions – i.e. performance would be worse during operations. The remaining factors may lead to an underestimate of operational performance in the test trial. Nevertheless, relative differences between baseline performance and performance with any COMDAT upgrades should become apparent.

3.6.9.2 Generalisability

The two major issues of generalisability are: (i) to other OROs beyond the T&E sample and (ii) to operational contexts. With respect to the former, as long as the required number of T&E participants can be obtained and given the small size of the overall ORO population, then adequate generalisability should result. Certainly generalisability will be much higher than in typical research endeavours where the test sample represents a very small proportion of the underlying population.

Generalisability to the operational context is subject to the issues outlined above with respect to operational realism. Increasing such operational generalisability may be addressed by an incremental approach to T&E in which the scenario may be repeated under increasingly higher levels of realism and involving greater contextual complexity. Thus, while initial evaluation may be conducted in NCOT, subsequent trials in the ORTT and then at sea would serve to provide data that confirm or disconfirm the operational generality of any findings. At present the generalisability of the data to other mission types such as assistance of civilian authorities and drug and smuggling interdiction cannot be estimated. By the same token it is not clear that MSDF technology, at least as exemplified by the initial TDP, is designed to assist decision making in these other mission contexts.

3.6.9.3 Risks: Implementation of MSDF Technology

The greatest risk that can be anticipated at the present time concerns how the MSDF TDP will be integrated into the existing simulation environments. We believe that Lockheed Martin intends initially to integrate the technology into the CSTC environment. As we have indicated previously, there appear to be significant limitations in the CSTC with respect to supporting the additional requirements for T&E, and collecting data with the required reliability and precision. Many of the planned measures outlined above would require a significant increase in capability of the CSTC simulation software to capture the required data. Even if the environment could be adapted to better accommodate the needs of T&E and the data could be captured appropriately, we would be faced with the situation of comparing baseline data collected in NCOT with MSDF trial data collected in the CSTC. Significant differences in a variety of variables across these two environments may either make comparisons impossible to interpret or reduce the sensitivity of the design to capture performance differences of interest.

A pragmatic approach to solving this problem might be to observe how MSDF impacts upon the information processed by the RT1 in terms of radar and track data, and then simulate these effects in NCOT for those T&E trials designed to evaluate the effects of MSDF. The practicality and feasibility of this approach of course remains to be evaluated.

3.6.9.4 Risks: the Need to Gain Direct Familiarity with Operational Functions

The accumulated knowledge of the T&E team to date has been based upon training exercises in the CSTC, evaluation of the Ops Room deficiencies, several days of scenario based interviews with OROs and other command team members, and information provided by Navy SMEs within the team. The MOPs recommended above are in many cases for behaviours that have not been closely observed but only inferred from verbal descriptions. Clearly what is required to round out the knowledge of the team is an opportunity for direct, sustained observation of actual Ops Room functions in progress by the T&E team. This will mitigate the risk of pursuing inappropriate MOPs and allow the T&E team to make a more informed final decision on which MOPs to include in testing

Now that the ORTT is in place and fully functional, we can believe that this can be readily and easily achieved. Therefore, *we recommend that before conducting the pilot trial, the T&E team either visit the ORTT while a full training exercise is in progress, or review existing records of training using the ORTT playback capability*. The goal of this will be to observe each of the major Ops Room functions in execution, with a view to validating the proposed MOPs and possibly uncovering critical tasks that may have been overlooked, for which MOPs should be developed. Of the two options proposed, the playback of existing scenarios may have several advantages.

First, the T&E team would not "get in the way" of live exercises or simulator training. Second, the team can stop and replay the scenario for better analysis. Third, a Navy SME can be used to assist in the analysis without any of the time pressure that occurs in live scenarios. Fourth, it may be possible to review records of several different teams to determine the degree to which processes are standardised or vary.

The logistics of this would involve three members of the T&E team (including one Navy SME) conducting a three day visit to the ORTT. If the playback review of previous records option is chosen then the analyses would be conducted at times when the main ORTT simulation facility is not being used by the Navy.

3.6.9.5 Risks: Lack of Data on Using Training Simulators to Measure Operational Functions for T&E Purposes

Worldwide, there is very limited experience in using training simulators for T&E purposes, and human performance data from operations or simulators is limited. The T&E demands for precision, reliability and repeatability may not have been the anticipated in the design of Navy simulation suites whose first priority is training. This is evidenced by the difficulty encountered in trying to retrofit measurement technology in the CSTC and the resulting insufficiency of the data produced to meet the more rigorous needs of T&E. Further, in the case of the NCOT facility, the proposed T&E program may push its envelop of capabilities and will be the first in-depth attempt to run team-interactive scenarios with multiple interacting workstations.

3.6.9.6 Risks: Availability of OROs as Test Subjects

The test plan is predicated on the assumption that sufficient OROs will be available to allow multiple test sessions in order to gather reliable data and establish confidence in estimating effects of interest. However, the reality that must be faced is that potentially fewer OROs could be made available to meet the needs of the program. If this were the case, then an appropriate approach would be to collect more data (in terms of replication of events) for each ORO. To accomplish this an additional scenario would need to be built with sufficient variation in events and actions to reduce carry over effects and minimise the risk of anticipatory responses. The project plan of resource requirements does not presently include this potential requirement.

3.7 Summary

The major tasks outlined above have been to identify appropriate MOPs, to consider how they may be implemented in a T&E scenario, to construct an initial overall T&E plan, to consider test design implications and to specify logistical and resource requirements for testing.

In reviewing the MOPs selected, it should be noted that the focus has been largely on those functions that impact upon the ORO's situation awareness process of detection, integration and comprehension of information from various aspects of the tactical pictures to support command

decision making. As such, these areas stand to be most impacted by the short to medium term MSDF technology developments of COMDAT1. For many of the other principle functions of the ORO concerning people and process management, the shape and potential viability of the MSDF technology that may enhance such processes is unknown. Therefore, no attempt has been made to consider appropriate MOPs for these functions at this stage.

A large number of MOPs have been identified, probably more than can be practically implemented in the main T&E trials to collect baseline data. Those measures that turn out to be impractical, unreliable or requiring undue overhead for the return, will probably need to be set aside, as lessons are learned in proof of concept and pilot trials. However, a further consideration that must be taken into account when evaluating whether to retain an MOP concerns its diagnosticity or overall effectiveness. The fact that a specific MOP may be accurately and reliably measured does not address the issue of its overall utility. For example, if a particular sub-process in the detect-to-classify sequence can be accurately measured, little useful information will have been gained if it is found to contribute to say less than 5% of the time required for the overall function. A priori, in the absence of detailed information flow process diagrams for these tasks with associated network simulations of process times, or being given access to operational performance data sets such as those collected by the Maritime Warfare Centre, we cannot provide guidance as to which MOPs are likely to account for the majority of the variance associated with effectiveness. Thus, it appears likely that as data are collected during the pilot and early main trials, evidence will accumulate as to which key MOPs will become the focus for optimum human-system performance description and analysis.

Finally, it should be noted that the CTA and subsequent validation provided an initial overview and framework for understanding the work of the ORO based upon a particular scenario structure. As we continue to understand the role of the ORO under a variety of operational contexts in real time, there will probably be a need to refine, augment and update the original analysis. As will also be the case to reflect ongoing changes in Navy command and control concepts, terminology and Ops Room resourcing strategies.

3.8 References

1. Matthews, M.L., Webb, R.D.G. and Bryant, D. Cognitive Task Analysis of the Halifax Class Operations Room Officer. DCIEM Report # . March 1999.
2. Matthews, M.L. and Webb, R.D.G. Validation of the Cognitive Task Analysis of the Halifax Class Operations Room Officer. DCIEM Report # 2000.
3. Matthews, M.L., Webb, R.D.G. and McCann, C. A Framework for of Military Command and Control Systems. DCIEM Report#CR-1997-024.
4. COMDAT I Technology Demonstrator. Statement of Work. DREA 2001.
5. COMDAT I Technology Demonstrator. System Concept. DREA 2001
6. The Naval Command and Control Way Ahead: A Blueprint for 2010.
7. AUS-CAN-NZ-UK-US (1997).Handbook 5: Guidelines for Maritime Information Management (version 2.1).Command, Control and Communications Committee.

THIS PAGE INTENTIONALLY LEFT BLANK

4. Technical Memorandum:

FINDINGS OF THE TEST AND EVALUATION PROOF OF CONCEPT TRIAL AT THE NCOT FACILITY (TRIAL DATE MARCH 11-13, 2002)

4.1 Summary

4.1.1 Background

Following the recommendations an initial assessment of NCOT in the previous report, this Technical Memorandum reported an assessment of the suitability of NCOT to provide the required environment for T&E trials to gather human-system data on small Ops Room teams with respect to:

- Scenario building, testing, evaluation and suitability;
- General logistics for time required, personnel support, robustness, audio/video recording, communications and fidelity;
- T&E suitability with respect to real-time control, communications between actors and subjects, consistency of scenario repetition, data recording and playback and analysis.

The assessment was conducted over three days. The focus of the assessment was a small sub-team comprising an ORO, SWC, TrackSup and RT1 with other critical roles being provided by the T&E team.

4.1.2 Findings

4.1.2.1 Scenario

Since there were no pre-existing NCOT scenarios that were appropriate for the T&E program, some initial scenario development was required. The scenario comprised background events in the air and on the surface from commercial and other traffic. Superimposed upon the background scenario were a number of potential threat events arising out of the specific geo-political circumstances. The scenario was initially planned on paper then coded into NCOT.

It was found that scenario playback in real-time was satisfactory; while some errors were observed in the characteristics and behaviours of entities, these would be easily rectified by altering the relevant database entry. Mission foreground events were suitable and could be discerned by participants. Workload for the scenario was not high, but this could be remedied through appropriate changes to the scenario, including adding events that overlap in time and including background tasks that would normally be undertaken.

4.1.2.2 General Logistics

The time required to set up and prepare a scenario to run included a full day to train SME actors, 30 minutes initial set-up, a 30 minute daily software set-up, 30 – 60 minutes daily scenario set-up, a 30 minute briefing period, a 30 minute 'mini-watch' (in order for participants to build 'the picture'), and a 10 minute watch handover. This means that each data collection day must include two and a half hours for these housekeeping tasks (in addition to the full day for training SME actors).

It was determined that a total of six people, plus those required for training and role playing, would be required to run an NCOT scenario for data collection purposes. These people were: NCOT Technical Support (on site), Trial Coordinator, Data Driver (1 & 2), Information Provider, ORO Observer, and appropriately qualified Navy SMEs for technical training support and role playing.

During the Proof-Of-Concept (POC) NCOT froze several times due to software problems. The 'Sleep' function also engaged periodically, necessitating a restart. This could be overcome by saving scenarios in 30 minute segments. Deficiencies were also observed in the simulated SG150 and STIR (Self-Targeting Illuminating Radar) and in the accuracy of the maps. Audio and video recording using equipment external to the NCOT system appeared to adequately capture data, and communications between participants using comms networks was also adequate, although often on the wrong network.

4.1.2.3 T&E Suitability

Real-time control over the movement of entities in the scenario was found to be possible but difficult and not necessarily realistic. Use of the predetermined track function would most likely be a better option. Communications between the actors and the Ops Room team were generally good, except in the case of the actor being engaged in other tasks. The scenario played consistently but (as alluded to above) there was some difficult with the consistency of dynamically-controlled events. Overall, NCOT seems capable of supporting a range of techniques necessary to capture MOPS, although automated MOP capture ability was not determined.

The recording of trial data is done on proprietary software and therefore not exportable to other platforms (e.g. other PCs). File management is rudimentary and will overwrite existing files if the hard drive is full. Playback is limited for a variety of reasons: each workstation can only play back its own record; play back can only occur in real-time; play back sometimes freezes if there are multiple concurrent play backs; and the play back audio can sometimes be obscured by background noise. In contrast, video play back is good. From the audio and video record it is possible to extract some MOP data.

4.1.3 Conclusions

NCOT provides an adequate environment for small scale scenario building and the collection of MOPs using small sub-teams. However, there are some serious limitations with respect to scenario play back for analysis and data and scenario record file management. The report has a number of recommendations for improving logistical aspects of conducting trails in NCOT in order to extract reliable MOP data.

4.2 Introduction

This work continues an ongoing program of test and evaluation (T&E) whose goal is to create reliable and valid measures of ORO performance in the Halifax class Ops Room, with a view to assessing the impact of future decision support technologies.

This Technical Memorandum summarises the outcome of the Test and Evaluation (T&E) Proof of Concept (POC) Trial at the NCOT facility that was conducted in accordance with the T&E trial plan outlined in reference 1. The POC was conducted by HSI ® team members during a visit to NCOT on March 11-13 2002.

This memorandum is organized into the following sections:

- Goals of the POC
- Method used
- Results of the assessment
- Conclusions
- Recommendations

4.3 Goals of the Proof of Concept

The overall objective was to assess the suitability of the NCOT facility to provide the required environment for future T&E trials to gather human-system performance data on small Ops Room teams. The intent of the trials would be to generate reliable performance data on a variety of C2 tasks using a simulation of the existing Ops Room equipment. A successful outcome would serve two goals. First, it would demonstrate that meaningful and reliable C2 MOPs could indeed be collected. Second, the data obtained would provide a basis for assessing the future impact of new COMDAT technologies.

The specific goals of the POC were to investigate the suitability of NCOT to support T&E trials in the following areas.

Scenario

- Building the scenario
- Testing that the scenario unfolded in real-time in the manner intended
- Evaluating whether the foreground threat events were suitable
- Assessing the suitability of the scenario for producing the required level of ORO and team workload

General logistics for running the scenario

- Time required for set-up and preparation
- Personnel support: NCOT technical support, Navy SMEs, T&E team (HSI®).
- Robustness and reliability of the system hardware and software during the scenario
- Ability to record additional audio and video
- Suitability and adequacy of communications among scenario players
- Suitability and adequacy of data supplied to scenario players
- Personnel support requirements:

Ability to support specific T&E requirements

- Control of game entities in real time
- Communications from T&E/actors to Ops Room Team
- Consistency of scenario repetition across trials
- Recording of trial data
- Playback and analysis of NCOT data files

4.4 Method used

A scenario was prepared ahead of time, details of which are provided in Annex C. The scenario was designed to create a moderate level of workload for the team and comprised background surface and air traffic overlaid with surface and air threat events. The Ops Room team required to run the scenario comprised an RT1, Track Sup, SWC and ORO. HSI® staff (ex navy ORO SMEs) provided additional information to the team that would normally come from the CO, OOW, CANEWS and any other relevant sources. The NCOT facility manager provided full time support to the team in setting up, running and analysing the scenario.

The structure of the POC trial was as follows:

Day 1:

- Play and review scenario in real time and correct any problems
- Configure workstations for the Ops room team
- Set up T&E logistics for audio and video recording
- Rehearse roles and control of game entities
- Review and rehearse requirements from NCOT support staff

Day 2:

- Brief Navy SMEs participating as subjects on purpose of POC, project.
- Review mission background for scenario with Navy participants.
- Familiarise Navy participants with Workstations and Comms
- Create necessary scenario overlays by RT1 and Track Sup
  (e.g. fixed reference points, area of patrol etc).
- Brief watch handover
- First trial run with scenario with interruptions to assess data analysis methodology
- Second trial run
- Third trial run
- Debrief participants
- Review trial playback
- Analyse outcome

Day 3:

- Brief new Navy SME participants
- Create necessary scenario overlays by RT1 and Track Sup
  (e.g. fixed reference points, area of patrol etc).
- Conduct watch handover briefing
- First trial run with ASCACT air profile patterns
- Break
- Second trial run with ASCACT air profile patterns

- Debrief team/washup
- Analyse/review trial runs

4.5 Results[17]

4.5.1 Scenario

4.5.1.1 Scenario development

The scenario was created from scratch as there were no suitable, pre-existing NCOT scenarios available. The scenario was based on a CSTC training scenario (OP CRATER) and was prepared using the existing game entities that were in the NCOT database, modified, where necessary, for present requirements. The scenario was first planned on paper with an outline of the geographical, political and military contexts, to which was added a Master Scenario Events List (MSEL). The latter provided a time ordered description of all of the game entities that would be introduced during the scenario and their associated trajectories. The full scenario description, briefing notes, order of battle, ID criteria, ROE and MSEL are provided in Annex C. The scenario was created by HSI® staff (ex ORO and Sea Trainer) with assistance from the NCOT manager. The scenario was intended to play for approximately 2.5 hours. The resulting master scenario file was stored on the NCOT system for later playback during the POC trial.

4.5.1.2 Did the scenario unfold in real-time in the manner intended?

In general, the playback of the scenario in real time was found to be satisfactory. Entities moved in the manner planned and appeared to be realistic to the Ops Room team. Some dynamic entities became de-activated unexpectedly during the course of the playback but this did not appear to have any apparent consequences to the players. The reason for these de-activations was unclear. They could have been caused by operator error, but the nature of the de-activations suggests that they may have been caused by an error in the NCOT software. Some entities did not have appropriate characteristics, for example Mirage aircraft were inappropriately provided with IFF. However, such problems can be readily rectified by correcting the appropriate database entry.

4.5.1.3 Were the mission foreground threat events suitable?

Air and surface threat events moved in the scenario appropriately and appeared to produce the appropriate responses among the team. That is, they were distinguished from the background events and initiated the expected level of assessment and analysis.

4.5.1.4 Was the scenario suitable for producing the required level of workload?

The team seemed to be operating at a moderate level of workload. They were able to keep up with the rate of events and did not appear to be dropping tasks. It should be noted that there were no overlapping air and surface threat events. In order to increase the workload level, the pace of air threats could be increased and overlapping temporal threats could be implemented. Further, there is a need to create some realistic level of background tasks that the ORO would normally face, such as managing the ship's Flex program or helo flying program and dealing with communications. This

[17] In the subsequent text recommendations are highlighted with italics.

will ensure that the ORO will need to divide attention over the normal task range for an ORO rather than to able to focus only on selected aspects of the tactical picture.

4.5.2 General logistics for running the scenario

4.5.2.1 Time required for set-up and preparation

This can be broken down into the following components and excludes scenario preparation, which was covered earlier. Note that the first and second two items can be done in parallel

Initial training day for T&E SME actors

Although not required for the POC trial, for all subsequent trials there will be a need to set aside time for training the Navy SME's who will serve as actors/confederates while the scenario is played out. They will be given an introduction to the scenario and their roles, and will participate in a full rehearsal of the scenario at least once, and possibly twice. During the rehearsal they will be trained to follow certain scripted requirements demanded by the T&E trial plan. Suitable breaks were inserted at approximately 2 hour intervals.

Initial physical setup - 30 minutes

This involves configuring the hardware, installing microphones and video recorder.

Daily software setup - 30 minutes

This was performed by the NCOT manager but, potentially, could be done by any suitable trained staff member. The above time allows for some re-starts as some problems were observed in some workstations that failed to be initiated or went down.

Daily scenario setup - 30-60 minutes

This involved having two navy SMEs (during the POC these were the RT1 and Track Sup players) manually enter fixed geographical entities into the CCS. This procedure had to be repeated every time the system went down. This step could be eliminated *if a provision were available to create scenario overlays that could be brought up at each SSD station.* If this cannot be achieved, then two people would be required to do this. In this study, the Navy SMEs acting as RT1 and Track Sup would be available to do this – but HSI® staff could also be used.

Briefing participants - 30 minutes

This includes, in this order, a study briefing(15 minutes), mission briefing (15 minutes) to familiarise study ORO subjects with study data capture procedures, and to commence building their "picture" or mental model of the mission and operational context to the point at which it would have been at the end of the watch preceding the start of the "study" watch, discussion of ROE and ORBAT issues with ORO subjects.

Mini-watch-30 minutes. This provides an opportunity for participants to become familiar with the specific hardware and to start to build the picture under light background load conditions. After the mission briefing, participants (and actors) go through a mini-watch (say 20 minutes), which would comprise the following elements. Starting a watch; doing a comms check; some light picture building (no surprises – and told that before they start this mini-watch); settling in, becoming familiar with the "picture"; building their mental model about what is going on; learning the sound of the voices of the different team members, how they request and receive information from sources not physically present in the scenario; getting used to the little NCOT peculiarities and room layout,

also any study specials (text message handouts, stateboard substitutes, pauses, SitReps, etc). In general the goal is to achieve the necessary familiarity so that they would not be experiencing these for the first time in the study watch proper. Also, during this period they get to ask any questions they want about the way things are going to work.

As a result, when they start the study watch, it will be more as if they are coming back on watch to something familiar, but just shifted forward in time.

Watch Handover (10 minutes) : Standard mission briefing. ORO provides SitRep at end to demonstrate understanding

4.5.2.2 Personnel support

NCOT technical support

Support was provided by the NCOT manager who is knowledgeable in matters relating to the running of the NCOT software and issues relating to hardware. No local software engineering support is available to deal with issues relating to the Unix operating environment. Such support is currently provided through telephone to Richmond, B.C. Some issues concerning software capabilities (such as the size of scenario log files and the over-writing of scenario log files) arose that could not be satisfactorily answered.

We recommend that system technical support be provided on site during the initial setup day for future trials.

Navy SMEs: technical support and role players

Technical support

Normally, NCOT is supported by Navy training personnel familiar with functions of the individual Ops Room team roles and the NCOT system. During the POC trial none of these were available, thus if problems occurred with a specialized aspect of the CCS functionality there was no one at hand to check out the problem and provide a solution. An example of such a problem that had impact on the fidelity of the simulation was our inability to flash up fire control radar as expected on a contact in order to get altitude information. Also, there was uncertainty about the provision of EW bearing lines in the absence of a CANEWS work station and operator.

We recommend that the appropriate NCOT Navy training specialists be made available on the set-up day of future trials to ensure that the NCOT system is providing the appropriate Ops Room functionality for all systems to be implemented in the trial.

Role players

For the POC trial, experienced RT1 and Track Sup operators from CFNOS were provided and met the requirements for the trial. The ORO position was filled by an ASWC who is currently taking the ORO course and the SWC position was played by someone who had not performed that role in 10 years and only had operational experience as a SWC with a ADLIPS display, not a HALIFAX SSD.

This experience highlights the fact that *future HSI® requests for Navy personnel must be precise and provide clear requirements for the experience level required for each team player and for the Navy to concur – or put the study at risk.* Failure to obtain personnel with the appropriate experience will clearly have the potential to render any ensuing data as unrepresentative of true operational performance. It should be acknowledged that not having individuals with the

experience requested for the POC trial may have been due to the short 4 week lead time available to the Navy in the request for personnel. We need direction with respect to realistic lead times required, and a better understanding of likely sources of suitable navy personnel.

T&E support staff

The following positions were found to be required to support scenario implementation and data recording of the trial.

1. Trial Co-ordinator: watches closely the unfolding of the mission events on the CCS and provides real-time co-ordination and adjustment of the scenario events to the NCOT Instructor Stations and data capture among the three previous positions. Filled by HSI® staff with extended training. This position is also the overall trial co-ordinator and is responsible for managing all aspects of the trial on the day.

2. Data driver 1: This person controls the background air and surface tracks that must be inserted or manipulated in some way in real time. An NCOT Instructor Station would be used for this position. Can be filled by HSI® staff with minimal training.

3. Data driver 2: This person controls all of the air and surface threat tracks that must be inserted or manipulated in some way in real time. An NCOT Instructor Station would be used for this position. Can be filled by HSI® staff with extended training.

4. Information provider: This person responds to all communications from the team with respect to requests for information and simulates the following roles: TG, CO, CANEWS, ORS (and in the future ASWC and TAS Sup). An NCOT SSD station would be used for this position. Can be filled by HSI® staff with Navy Ops Room experience (e.g. ex navy ORO).

5. ORO observer: sits behind the ORO and observes/assesses ORO actions in real time. Requires SME with ORO training experience - be provided by HSI® team.

Comments.

Initially in the POC trial we tried to combine positions 3 (data driver) and 4 (information provider) into a single role. However, the workload proved to be too high for a single individual such that communications to and from the team were subject to operationally unrealistic delays or unintentionally inaccurate responses. Much of this workload resulted from EWS traffic – particularly related to the absence of automatic EW bearing lines from CANEWS.

Positions 1 and 2 are only required because of limitations in the existing NCOT software that currently require entities with changes in their trajectories to be entered into the scenario manually. Ideally, we would like to pre-script all events and have all entities enter into the scenario and follow their scripted trajectories automatically. Such a script is simple for background air and surface contacts that move more or less on a single trajectory across the area of interest. In the case of threat and other similar entities, whose trajectories may involve frequent changes in course, altitude or speed, the script could include all of these desired characteristics, such that the entities always move in a known prescribed pattern. Therefore, one major software improvement that would obviously reduce T&E personnel support would be to *provide additional software functionality that would allow entities to automatically enter the scenario and have their trajectories controlled by the NCOT software following a time-ordered script.* This requirement corresponds to the existing NCOT function of "**Predetermined Tracks**" which allows the creation of pre-determined waypoints for an entity to follow. The preparation work for the scenario revealed that this function was not operating in the expected manner. This is an important NCOT

function for the support of the T&E program since it allows the creation of tracks that can also follow a prescribed path such as zigzag, ellipse and sinuation as well as formation flying in which cohort entities automatically follow the pattern of the formation leader. The actual availability of this function needs to be checked as soon as possible.

4.5.2.3 Robustness & reliability of system hardware & software during play

The system malfunctioned several times during the course of the POC trial. The most frequent type of failure was a freezing of the scenario and workstation, which then required a re-boot of the entire system (i.e. all work stations). The malfunctions appeared to be mostly software induced, although in one case there was a hardware failure. Further, each workstation had a "sleep" function, such that in the absence of any keyboard or mouse activity for a period of time (the exact period was not known and differed for each workstation) the workstation would shut down and the entire network would require re-booting. This resulted in a time-out of approximately 10-15 minutes while the scenario was restarted. A further 10-20 minutes was also required to allow the RT1 and Track Sup to re-enter CCS data such as geographical points etc. Yet more time was required, since the scenario could not fast forward to the point of stoppage without the re-creation in real time of the threat and background tracks. This means having to replay parts of the scenario that the team will already have been through. *It is recommended that the "sleep" function either be switched off or adjusted to a much longer time-out interval.*

A possible work-around for this problem would be to save the scenario file in sequential fragments of 30 minute duration. Using the appropriate file segment following a stoppage would mean a worst case situation where 15 minutes of scenario would be repeated. The downside of this would be the time required to rebuild the scenario overlays and a loss of operational realism.

A second type of failure appeared to be related to the specific functionality of some of the simulated Halifax class systems. Such failures were noted as follows:

- SG150 radar: failed to auto-initiate and auto-track surface tracks, even when manually generated tracks were assigned to the radar
- STIR fire control radars: failed to respond to primary designation by SWC, thereby not providing altitude and fire control information on the contact

A third, and possibly less important problem, was found with the NCOT landmass database in the area of operations. Small islands north of the patrol area were not depicted and there was a westward exaggeration of about 6 miles of the Yemen coastline. While causing some reduction in operational realism, we believe that this problem is manageable

Our preliminary assessment is that such failures give rise to some concerns about whether the NCOT facility can provide the required level of reliability to support the T&E effort. A final judgment of this should probably be suspended until after the pilot trial has been conducted since some problems may simply require better insight into existing system functionality.

4.5.2.4 Ability to record additional audio and video

A video camera was mounted towards the rear of the team and appeared to adequately capture the movements of the team (e.g. head position, body movements, face to face interaction). Microphones were mounted between the ORO and SWC and between the RT1 and Track Sup; and these appeared to be adequate for capturing off-net verbal communications and verbal input to radio microphones from those positions.

4.5.2.5 Suitability and adequacy of communications among scenario players

The NCOT communication facility appeared to be adequate in providing the requisite network communications among the team, although some of the circuits were not able to be assigned their usual SHINCOMS reference. For example, the internal Above Water Warfare coordination net was assigned to the ASW button instead of the AWW button. The quality and speed of communications appeared to adequately mimic normal Ops Room circumstances. Recording quality was marred by noise from alarms, which blotted out some recordings. There was also a background hum from the ventilation system that made softly spoken communications difficult to hear.

4.5.2.6 Ability to support specific T&E requirements

Control of game entities in real time

Game entities are described in detail in Annex C. Once introduced into the scenario, game entities generally moved in an operationally realistic manner. There were three types of entity patterns that needed to be controlled - single trajectory tracks, tracks that had several legs and multiple trajectories and tracks that may need to respond or react in an appropriate manner in real time to actions taken by the Ops Room team.

Single trajectory entities are brought into the game at a known time and reference point and follow a ballistic pattern across the area of operations. Examples include, commercial air traffic taking off and landing , helicopter traffic between the coast and oil-rigs and surface vessels sailing from ports. The T&E team controlled these entities by taking them from a "parked" area that was outside the area of operations and moving them to a waypoint and activating them. At this point the scenario software then took over their movement. This manual control process was required because the current version of software does not support the timed entry of entities into the scenario following a pre-scripted schedule.

The second type of tracks involved multiple trajectories, that is entities that might change course, altitude or speed while in the area of operations. For example, Yemeni aircraft taking off from Hays airbase could follow a variety of patterns that might provide different levels of threat and interest. Each individual change of pattern required dynamic real-time control by a member of the T&E team.

This proved to be a manageable task, but was subject to a number of limitations. First, as the number of entities in a complex scenario that require real-time control increases, the ability of the entity controller to manage all of these becomes compromised. Second, there is some imprecision in following the required trajectory script exactly and to repeat that pattern in the same manner on successive trials. This results from small lags in initiating events and latencies in aircraft flight dynamics, as well as the skill level of the operator. Approximate patterns can be repeated with ease; however, this will result in some imprecision in collecting MOP response time data where the start time is initiated by a track that is supposed to be at an exact geographical location, relative to the ship. We were not able to establish the magnitude of such potential errors because of problems with scenario playback (see below).

The requirement for precise repeatability of tracks is particularly relevant to the generation of tracks that are designed to produce some radar confusion, as exemplified by the patterns used in ASCACT trial. Three of the ASCAT patterns were evaluated. Pattern 1: two aircraft at same altitude and speed, on converging tracks, converge and fly close together directly over the AOI

before diverging. Pattern 2: two aircraft at the same altitude and speed have trajectories that cross backwards and forwards with each other across the AOI. Pattern 3: two A/C on same heading and speed approach AOI, one A/C changes altitude lower, and maintains lower altitude while holding same direction as other A/C.

In general, it was found difficult to generate tracks involving two A/C doing such close manoeuvring with exact repeatable precision, although, with practice, a reasonable approximation to the required patterns could be produced. Further, some difficulty was encountered in getting entities to make fast turns in the manner that accurately simulates combat manoeuvres.

One important observation was that the NCOT simulated radar seemed to experience no difficulty in maintaining separate and correct track identities for the ASCACT tracks that we were able to reproduce, such as two A/C coming together from different vectors. Therefore, it remains a moot point whether to continue to use such tracks to simulate existing radar tracking problems that MSDF technology is designed to solve. Therefore, *we recommend an early discussion with DREA and LM to further elaborate upon the development of track requirements that will suitably address existing radar technology problems and MSDF improvements*. *We further recommend that DRDC put in place the necessary contractual arrangement to* allow HSI® *to interact and work closely with LM to ensure rapid and timely interchange of technical information concerning the progress of COMDAT support work.*

The third type of track, which requires the game entity to move in an operationally realistic manner in response to the appropriate circumstances in real time, would require real-time manual control by the T&E team in any case.

Our assessment is that in order to reduce T&E overhead for manual control of entities and to ensure precision in the execution of track trajectories *the software be enhanced to allow pre-scripted tracks to enter and move within the scenario under game software control.*[18]

Communications from T&E/actors to Ops Room Team

One member of the T&E team played multiple roles to provide contextual communication and information to the Ops Room team by way of one of the radio nets in response to requests for information from the CO, OOW, CANEWS etc. This appeared to work effectively and in a timely manner. The exception, noted earlier, was when this same member of the T&E team was concentrating on the control of game entities, which resulted in a lag in the communication response.

In future trials, it would appear possible and desirable to pass paper text messages to the ORO using the ORO observer position. These would make the ORO task more realistic, and include a requirement to select and forward relevant information to others in the Ops room team.

Consistency of scenario repetition across trials

Although we were not able to assess this formally because of playback problems, we were able to conclude that background air and surface events could be repeated with the required precision across trials. However, there were problems in maintaining the required consistency of both the timing and spatial precision of dynamically controlled tracks.

---

[18] We have been told that this was supposed to be achievable through the "Pre-Determined Track" function of the Scenario Builder software, however, we have not had any success in making this work correctly.

Support for collecting specific MOPs

The detailed MOPs provided in reference 1 were reviewed and a generic list of requirements to support the delivery or required scenario events and data collection was constructed. The generic capability was derived by examining all of the individual scenario input, monitoring and recording elements, and then grouping them into more generic categories. This information is shown in the Tables 4.1-4.5 below, together with an indication of the ability to support these requirements within NCOT and by the T&E team.

Table 4.1 provides a summary assessment from the POC of the capability of NCOT to provide information that must be provided to the ORO other than through the specific scenario events. This information could be text messages by hand or via the CCS, or voice messages through the various nets or face to face. The meaning of the various columns is as follows:

Info to ORO: describes the type of message and medium

Feasible: indicates our assessment of whether it can be done in NCOT

Timing sent: means if the time of the message transmission can be recorded and how this is accomplished.

Timing recd: means if the time of the message reception by the ORO can be recorded and how this is accomplished.

T&E-who does: indicates who in the T&E team would be responsible for delivering the message.

If not T&E, ID source: If the message does not originate with the T&E team, can the message source be identified and how.

Capture Response Action: If the message results in an ORO action (response) can this event be captured and how.

Capture Response Content: If the message results in an ORO action (response) can the content be captured and how.

ID target of response: If the message results in a response by the ORO can the target (recipient) of the response be identified.

Time of response: If the message results in a response by the ORO can the time of the response be captured.

Assess response: If the message results in a response by the ORO can the content of the response be analysed and how.

In Table 4.2 , we assess the outcome of the POC in terms of the capabilities of NCOT to provide support for the following input events

- Background air/surface commercial
- Background air/surface hostile
- Significant A/C actions (changes course to TG, altitude, speed, radar, previous pattern, assumes attack profile
- ASCACT aircraft patterns thought to generate radar confusion.

For each of these we determined that each of the following required pieces of data for measurement purposes could be obtained from the scenario or playback:

- time into scenario

- time painted on radar display
- time into area of interest
- time out of area of interest
- time of change ins status of contact
- time at own ship's weapons range
- time at A/C weapons range
- time EWS emitted
- time EWS detected.

While the majority of these input events can be easily implemented in NCOT and the required data captured, some technical difficulties in simulating ASCACT tracks to the required degree of precision was encountered.

In Table 4.3 we show how in NCOT we can capture data relevant to ORO actions on the CCS, and interactions with other information sources. Tables 4 and 5 show how we can record ORO and other communications and assess ORO knowledge, respectively. This capability is assessed in terms of how we would specifically capture information content, the time of the action and the identity of the recipient of ORO communications. The ability to add T&E embedded probes into the scenario is also assessed using approaches such as a SITREP to the CO, a "bogus" request for information as well as the means for determining the associated response time to such embedded probes.

In general, NCOT appears to be capable of supporting the range of techniques that will be necessary to capture MOPs. Having said that however, it should be noted that at this time we have not had the opportunity to actually calculate sample MOPs from the POC data record. This arose because of a number of problems relating to the lack of reliability in the system functioning and especially in playback (see below). Clearly, the pilot trial will provide the most opportune time for assessing the feasibility of capturing and analysing the data in the manner intended.

Recording of trial data

For each workstation the screen content and all net based communications are continuously recorded on the hard drive of the individual workstation. This information is stored in a proprietary file format and is not amenable to being read using standard data analysis or statistical software. The entire scenario session may be recorded as a single file, or recording can be started and stopped under real-time control to yield smaller, sequential data files. This may have some advantages given limitations in data playback outlined below.

However, file management functionality appeared rudimentary, and could be improved. For instance, we were not able to access information concerning the length of the resulting file. Whether this is a result of a software limitation or lack of knowledge on our part on how to retrieve such information is not known. We also learned that once a hard drive becomes full the system starts to overwrite existing files; again how this actually happens and the consequences for stored data from trials is unknown. At present files cannot be copied or removed from the hard drive onto secondary mass storage media.

These issues raise some concerns for the future management and access of T&E data. Hence*, we recommend that a capability be added to store data files on an external medium*. This may require a DVD format given that the length of the file may exceed CD-ROM capacity. A further advantage of storage on such a medium is outlined in the following section concerning playback of stored data. Whatever medium is chosen, this requirement is considered essential for the retention of critical trial data, the production of which, has been the result of considerable logistical effort, which, obviously, cannot afford to be wasted.

Playback and analysis of NCOT data files

Playback

Some serious limitations to the playback of NCOT data were discovered and are outlined below.

1. Data files are currently 'owned' and located on each individual workstation, as a consequence, they can only be played back on that station. The playback may be monitored on a separate workstation that is temporarily configured for this purpose on the NCOT network. The NCOT system does not produce a single data record integrated across all of the workstations that comprise a small team configuration for a particular trial. This lack of capability may have significant consequences for trial analysis. Using the NCOT playback, the simultaneous review and analysis of the T&E trial data across all team workstations would require a co-ordinated manual start of the replay on each of the individual workstations and an observer/analyst at each workstation. While, with some minimal practice, any error in ensuring a common start of the replay would probably be acceptable, we have not been able to ascertain whether the playback at each workstation will remain synchronised over what could be the course of a 2 hour trial.

2. Data files can only be played back in real time. Further, there are no capabilities to fast forward, slow forward, reverse, pause or stop/restart. In fact, once the replay has been stopped when it is re-started the playback reverts to the beginning of the file! Clearly, such limitations severely compromise the ability of the T&E team to perform the required fine grain analysis of the trial data in anything approaching an efficient manner, if at all. What is already likely to be a labour intensive activity is only made worse by this apparent lack of a rudimentary capability.

Two solutions to this problem are suggested. First, that the necessary software could be developed to provide the functionality to support analysis of trial data (outlined in detail later). This approach is likely to be time consuming and costly. *The second, and our preferred, solution is to provide a capability to capture the video and audio output of the workstation playback and convert into a format that will allow recording on standard digital video media.* The added advantage of such an approach is that trial data could then be taken off-site for analysis at HSI® premises, with significant savings in travel overhead, or the need for the purchase of an NCOT compatible workstation for data playback. Such capture could be accomplished through the purchase of a high quality video converter that would transform the native NCOT RGB video output into an NTSC video compatible format. The technical requirements and cost of such a solution are currently being investigated.

3. Several system freezes were encountered during the playback of recorded data from the DRDC POC sessions.[19] This may have been the result of trying to playback several files simultaneously on different workstations in order to accommodate our needs. On the final afternoon, the system crashed during playback and could not be re-started to allow further playback. Telephone enquiries to MDA headquarters in Richmond did not result in an immediate solution being found to the problem of playback freezing.

4. Playback of audio traffic captured from the comms nets was somewhat degraded by the intrusion of noise from background fans of the heating/ventilation system. Further, communications became almost unintelligible when a CCS alarm had been captured on the audio track. Our tentative conclusion is that the audio playback is of marginal quality for T&E analysis. Further investigation of this is required to determine whether there are ways to improve the signal to noise ratio.

On the positive side, the quality of the video playback, whether on a workstation monitor, or displayed on a large, front-surface projection screen through a high-resolution video projector, was sharp and provided clear detail of all symbology, alpha-numerics and QAB strokes. However, a separate video record, that was captured from a workstation monitor or the projected image by the T&E team digital video camera, showed poor image quality. Thus, it was concluded that this would not be a satisfactory workaround to the problem of creating a video record that would be more amenable for T&E analysis than the NCOT data file.

Analysis

Notwithstanding the above comments concerning limitations, the data files do provide the necessary information to allow detailed analysis to allow the extraction and calculation of performance data. The workstation clock is visible at all times and would allow timing of events (accurate to the nearest second) to be extracted. The effect of the operator's actions are clearly seen on the screen in terms of the information provided in the radar area as well as in the various numerical and tabular data areas. We were able to see which track was currently hooked, the track data, the range selected and all other information that provides the tactical picture. Thus, it would appear feasible to determine from the video and comms records the ongoing tactical focus of the ORO and also allow some response time measures to be collected, such as the time between the appearance of new information and the subsequent ORO response. It should be noted that because

[19] This did not happen as much with the earlier recorded DREA POC data that was stored as a single file. The DRDC data were stored in a sequential series of smaller files in order to avoid having to go back to the very start of the scenario, each time the playback was stopped.

of the lack of access to the playback files in the final session, we were not able to actually attempt to extract MOP data examples as had been planned.

Info to ORO (other than T&E software)	Feasible	Timing Sent	Timing Recd	T&E - who does?	If not T&E, ID source	Capture Response Action	Capture Response Content	ID target of response	Time of response	Assess Response
Text message via CCS	NO									
Text message by hand	YES: via ORO observer.	YES: ORO obs notes WS clock	YES: Trial coord notes appropriate WS clock	ORO obs Trial coord	Known to team	YES: NCOT video/audio record of ORO WS or ORO obs.	YES: NCOT video/audio record of ORO WS or ORO obs.	YES: NCOT video/audio record of ORO WS or ORO obs.	YES: from playback clock or observer record.	Post event analysis by SME
Voice message from net	YES	YES: video/ audio playback	YES: playback clock	NA	YES: video/audio playback	YES: NCOT video/audio record of ORO WS or ORO obs.	YES: NCOT video/audio record of ORO WS or ORO obs.	YES: NCOT video/audio record of ORO WS or ORO obs.	YES: from playback clock or observer record.	Post event analysis by SME
Voice message face to face	YES: from T&E video/ audio record	YES: need to add clock to T&E video record	YES:	NA	YES: from T&E video/audio record	YES: NCOT video/audio record of ORO WS or ORO obs.	YES: NCOT video/audio record of ORO WS or ORO obs.	YES: NCOT video/audio record of ORO WS or ORO obs.	YES: from playback clock or observer record.	Post event analysis by SME

Table 4.1: Feasibility and methods for providing information to ORO and for capturing MOPs

Input events from scenario	Time into scenario	Time painted	Time into AOI	Time symbology	Time Out of AOI	Time change in status	Time at own ship weapons range	Time at A/C weapons range	Time EWS emitted	Time EWS detected
Background air/surface-commercial Background surface Hostile air Hostile surface A/C changes course to TG A/C changes speed A/C changes alt A/C radar A/C changes previous pattern A/C assumes attack profile Neutral-changes profile ASCACT patterns	YES: from ground truth playback of scenario	YES: from ground truth playback of scenario	YES: from ground truth playback of scenario	YES: from playback of trial scenario	YES: from ground truth playback of scenario	YES: from ground truth playback of scenario	YES: from ground truth playback of scenario	YES: from ground truth playback of scenario	YES: from ground truth playback of scenario	YES: from ground truth playback of scenario

Table 4.2: Feasibility and method of implementing and recording scenario events

ORO actions	Capture screen	Capture/ timestamp action	Assess action
ORO action on CCS	YES. NCOT playback.	YES. NCOT playback. Accurate to nearest second.	ORO obs in real time commentary through headset. Post trial analysis by SME on NCOT/T&E playback.
ORO consults other sources	ORO use of ancillary information (e.g. manuals, plot, simulated stateboards, tacpacs) captured by T&E video and analysed post event. May also be captured by ORO obs.		

Table 4.3: Feasibility and methods for capturing ORO actions

ORO Communications	Capture message content	Capture/ timestamp action	ID recipient	Assess content
ORO speaks directly to team/member (not on net)	YES: T&E audio capture.	YES. T&E audio file playback. Will need to add functionality to provide visual time base indicator.	YES. T&E audio file playback.	SME analysis of T&E audio file.
ORO uses net	YES: NCOT audio capture	YES: NCOT audio playback plus CCS clock (video record).	YES: NCOT audio playback	SME analysis of T&E audio file.

Table 4.4: Feasibility and methods for capturing ORO communications

Assessing ORO knowledge	Real time SITREP to CO	ORO CCS Display	T&E Bogus request for info	Response time to embedded probe
Picture content at start of scenario	YES		YES	NA
Change in picture content	YES		YES	NCOT playback
Indiv threat comprehension	YES		YES	NA
Change in threat status	YES	YES	YES	NCOT playback
Relative threat priorities	YES	YES	YES	
Tactical situation	YES		YES	
Time to CPA	YES	YES	YES	
Own engagement envelope: earliest/latest point to shoot -		YES		
Enemy engagement range – contact		YES		

Table 4.5: Feasibility and methods for assessing ORO knowledge.

4.6 Conclusions

The NCOT environment provides *adequate support for the development and running of scenarios* to support T&E testing and for the collection of required data. This assessment is based upon the assumption that problems with the fidelity of simulation of some system functions may be readily corrected.

With respect to *the playback and analysis of T&E data we believe that there are some serious limitations of the NCOT system*. This conclusion is based upon the way we expect that the data will have to be analysed after the trial conclusion. We anticipate that we will need to have a high level of control over the scenario playback process in order to extract the required data; such control will include an ability to pause, rewind, slow forward and fast forward. Further, there may be a need to provide a time-synchronised playback between two or more workstations. In order to achieve such control, significant software changes would need to be made or alternate approaches should be considered (as outlined below).

A second concern relates to the ability to manage archived scenario files. At present we have been advised that there is no capability to determine the length of files or how much space is available on the drive to accommodate new files or to ensure that they are not overwritten. There is also no ability to copy data files to external devices at individual workstations.

A third concern is that, even assuming that the above two problem areas can be corrected, all analysis would have to be done at the NCOT facility. Given that one might expect each hour of recording to require 3-4 hours of analysis, this would mean the T&E analysts would have to spend several days at NCOT. This could not only raise issues of accessibility, but also add overhead costs to the project associated with travel and accommodation. Again, some solutions to this issue are outlined below.

4.7 Recommendations

4.7.1 Software and system improvements

(in approximate order of priority)

Some of these improvements may indeed exist, but were not apparent or could not be made to function during the POC.

1. Provide a capability in the master scenario file to allow entities to follow a pre-scripted trajectory. This could be readily accomplished by fixing the problems in implementing the NCOT Scenario Builder function of "**Pre-determined tracks**". (NCOT Manual Section 6-27).
2. Allow entities to automatically enter into the scenario at pre-specified times and follow a constant track and speed.
3. Allow entities to automatically enter into the scenario at pre-specified times and follow a pre-scripted trajectory.
4. Provide a capability for a time-stamped flag to be entered from an instructor workstation into the scenario record. This could be achieved by capturing function key presses and logging the time of the press and saving to a file. The resulting data file should comprise a time-ordered list of the key presses and time (scenario time) when

pressed. This file should be in an ASCII text/tab delimited format to allow for export to a spreadsheet.

5. Provide a capability to capture specific key presses at the ORO and other team workstations. These are to be time stamped and recorded to a data file in the manner described under item 4. Specific actions required to be captured include: quick action buttons (QABs), numeric keys, "enter" key, F2 and the alarm button.
6. Develop a capability for control over playback files to include: slow motion, pause/re-start, rewind, fast rewind/fast forward and go to flags (as in#4).
7. Provide an ability to download data files to external media.
8. Provide an ability to replay multiple files simultaneously on different workstations (preferably in synchrony)
9. Provide data file management capabilities
10. Provide an override of the workstation "sleep" function or allow it be adjusted to a much longer time out.
11. On SSD enable dimming of symbology so that can operator can examine radar paint alone (as occurs on the actual CCS)
12. Provide a capability for the creation of CCS overlays ahead of time and to allow them to be integrated into the Tactical Situation Area of the SSD when the scenario is run.
13. Enable instructor to draw "lines" and save for later use e.g. patrol areas (e.g. like Word Draw options plus text insertions)
14. Provide a means of distinguishing between active tracks and stationary or parked tracks in the instructor screen i.e. to rapidly identify all tracks in play at the moment
15. Allow audio playback on any workstation
16. Improve signal to noise ratio (sound quality) of audio playback.

4.7.2 Data capture and analysis

Notwithstanding items 3,4 and 5 above, our preferred solution for data analysis is to allow data files to be captured in a manner that will allow their removal from site and analysis to be conducted at an HSI® site. To achieve this we propose the following solutions:

Capture the NCOT workstation video output to the CCS display in real time and convert it to NTSC video using a scan converter. This will be required for each workstation of interest (which for present purposes is 4). The subsequent video signal can then be captured on digital videotape with a single record for each source. Possibly, if we can maintain signal quality, the four records could be multiplexed onto a single tape. For audio records, NCOT maintains a separate digital audio file for each workstation, and these could be subsequently transferred to a removable media to allow integration with the video record when the analysis is performed. We are also exploring a new technology that will simultaneously capture the output of four workstations and integrate them into a single digital (non-video) record for subsequent payback.

An alternative approach would be to purchase and configure an NCOT compatible workstation for local playback/review purposes. The major disadvantage of such an approach is that the playback

would be limited to a single data record at a time. Hence, the co-ordinated analysis of information across team players at different workstations could not be achieved.

4.7.3 Personnel requirements to support T&E

1. On the set-up day prior to data capture, NCOT Navy trainers be made available to ensure that all of the Ops Room simulated functionality is working in the operationally correct manner.
2. This day will also serve as a training day for T&E SME actors. They will be given an introduction to the scenario and their roles, and will participate in a full rehearsal of the scenario at least once, and possibly twice. During the rehearsal they will be trained to follow certain scripted requirements demanded by the T&E trial plan.
3. On the same day, a systems software specialist should be on hand to ensure that the system is functioning appropriately and that there is space to store data records.
4. *Navy provided SMEs*.
   Actors: 1 RT1, 1 Track Sup, 1 SWC. All need to have some experience, preferably including working together as a single team. Thus, our ideal would be to draw an intact team from one watch on one ship to stay over both days. Recognizing that this is unlikely, then we need SMEs that have relevant qualifications and relevant experience in their recent past, e.g. QL5 NCIOP for the TS. A QL4/5 NCIOP (or a QL3 NCIOP, with significant experience) for the RT1. Equivalent needs apply to the SWC SME. Whoever is provided, they need to stay for the duration of the trial (two days for the pilot, five days for the study) so that we don't have to repeat the start up briefing / learning time on subsequent days.
   Subjects: For the pilot study, two OROs will be needed for a half day each. For the actual study, a further eight OROs (each for half a day) must have completed the ORO course and have had at least one year as an ORO in a HALIFAX class frigate, as recently as possible, but certainly within the last four years.
5. *HSI® personnel*: The make-up of the T&E team for the Pilot Trial is largely dependent on whether software fixes 6.1.1 and 6.1.2 are in place. The following table outlines the team required, depending whether such fixes are done or not.

T&E function	Personnel Required	Software fixed	Software not fixed
Entity controller-background	HSI® staff - trained	NO	YES
Entity controller-threat	HSI® staff - trained	NO	YES
Information provider	HSI® - Navy SME	YES	YES
ORO Observer	HSI® -Navy SME	YES	YES
Trial co-ordinator (also controls some dynamic entities	HSI® senior staff	YES (also controls some dynamic entities)	YES
TOTAL		**3**	**5**

4.7.4 Scenario modification

To ensure appropriate distribution of ORO attention, some additional tasks will need to be provided. We are working out details of these and how they may be practically implemented.

4.7.5 Other issues

We recommend an early discussion with DREA and LMC to further elaborate upon the specific requirements for the development of "test" tracks that will address existing radar technology problems in track detection and maintenance and other track contexts that are likely to be benefited by MSDF improvements.

4.8 References

1. Matthews, M.L., Webb R.D.G and Keeble, A. R. Assessing the Impact of Multi-Sensor Data Fusion on Command and Control Operations in the HALIFAX Class Frigate: Recommendations for Measures of Performance and Detailed Test Plan. Section three, this report.
2. Naval Combat Operator Trainer, Volume 2: General Purpose Reconfigurable Trainer– Technical Description And Operation. 2001-08-20

5. Technical Memorandum:

EVALUATION OF THE RGB SPECTRUM DGX DIGITAL GRAPHICS RECORDING SYSTEM AS A MEANS OF COLLECTING NCOT MOP DATA TO SUPPORT COMDAT

5.1 Summary

5.1.1 Background

This Technical Memorandum was initiated to investigate the DGx digital graphics recording system as a potential remedy to the deficiencies identified during the POC. In particular, the DGx recorder might address improved playback capability to better support the T&E team's requirement to efficiently extract MOPs from the data record.

5.1.2 Findings

The DGx recorder offers a number of features that would be beneficial to gathering MOPs. They include:

- High resolution video capture from up to four NCOT workstations;
- Appropriate capture rates;
- Audio capture;
- Easy access to and replay of the data record;
- Off-site data access.

The demonstration was arranged for HSI® and MDA at the NCOT facility. A total of nine assessment criteria with 17 sub-criteria were used to assess the DGx recorder.

The DGx recorder needs to be set up close to the workstation that it is recording. A high quality video splitter is required for each workstation to be monitored. The DGx recorder only records the left and right audio (i.e. two channels), which is not sufficient to record a small ops room team. The total set-up time for the DGx recorder is between 30 and 45 minutes; there is very little intrusion on the needs of T&E simulation itself.

- Video quality – there are some issues (reading track numbers, jagged diagonal lines) but video quality was generally considered good enough for collecting MOPs.
- Frame rate – drops to 1.5 hz in high resolution, quad (four workstation) mode, but was generally considered satisfactory.
- Audio capture – no evaluation was made.
- Ability to set flags – good and easy to achieve.
- Playback – good, allows fast forward, rewind and stop. However, if the playback is stopped, the playback returns to the beginning. Remote playback must be on a dedicated monitor through the DGx processor.

* Data management – using hard drive storage a total of 9 hours can be stored; using tape storage a total of 3 hours can be stored. The combined capability afforded by hard drive and tape storage appears to meet the needs of T&E.
* Portability – the DGx recorder weighs 21 lbs and measures 18”x17”x3.5”. The DGx recorder consists of two units: the processor and the storage unit. Both units seem robust and therefore portable if necessary.
* Software functionality – the software operates on a PC running MS Windows, using a standard interface. This makes the system intuitive to use. There is no anticipated requirement for specialised software.

The total system cost would be $41356. This cost includes the DGx recorder and replay unit, a 3 hour tape data recording unit, a 120 Gb storage device, and the virtual control panel software. An additional outlay of approximately $2000 would be required for high quality video splitters.

5.1.3 Conclusions

The conclusion from the assessment was that the DGx recorder is a viable supplementary solution for collecting and analysing MOPs in NCOT and could also be used to capture data from a variety of other sources (e.g. ‘live’ situations). It was recommended that the means by which the system (and video splitters) could be purchased be investigated.

One possible alternative to the DGx recorder is a video scan converter and associated storage. This would cost approximately $10000 but one would be required for each workstation, resulting in a total cost of $40000 (i.e. roughly equivalent). However, this alternative would not provide easy synchronisation of the different recordings.

5.2 Background

A recent proof of concept evaluation has been conducted at the NCOT facility at CFNOS, Halifax to determine test and evaluation (T&E) capabilities for the collection of performance data in support of the COMDAT program (reference 1). One of the major conclusions was that there were some significant limitations in the NCOT playback capability that would severely impact upon the ability of the T&E team to efficiently extract and quantify measures of performance (MOPs). As a result, alternate methods for capturing NCOT data were explored.

After evaluating many options, including scan converters to allow the NCOT data to be recorded to a video medium, it became apparent that the DGx digital data recording system manufactured by RGB Spectrum seemed to offer several strong features which include:

* High resolution video capture from up to 4 NCOT workstations
* Capture rates adequate to the dynamics of the Ops Room
* Audio capture
* A data record that can be easily replayed and accessed
* Accessibility of captured data off site from NCOT

To further evaluate the suitability of the DGx system, a demonstration was arranged with the DGx Canadian Distributor (Integrys Ltd.) at the NCOT facility with the co-operation of MDA and CFNOS/NCOT support personnel on August 22, 2002; Michael Matthews and Roy Keeble were present from HSI and the NCOT manager from MDA.

5.3 Evaluation criteria

The following criteria were used to evaluate the system.

1. Ease and time to set up
2. Quality of video image capture, including the following elements:
   2.1 Alphanumerics in radar (e.g. track numbers) and data areas
   2.2 NTDS symbology
   2.3 Graphical elements such as lines and reference points
   2.4 Individual radar paints
   2.5 History trails
   2.6 QAB labels
   2.7 QAB selections
   2.8 Hooked tracks
   2.9 Colours (white, yellow, green, cyan, red, magenta)
3. Adequacy of frame rate
4. Quality of audio image capture
5. Ability to set flags to mark events during recording
6. Control over playback
   6.1. Four workstations in synchrony
   6.2. Stop-restart
   6.3. Fast forward/rewind
   6.4. Ability to set flags to mark events during replay
   6.5. Ability to go to flagged events
   6.6. Off -site capability
7. Data management
   7.1. Data storage/archive
   7.2. Data format
8. Portability
9. Software functionality

5.4 The Assessment

Because the DGx has a number of recording modes, some preliminary decisions must be made with respect to the number of screens to be captured and the desired resolution. This decision influences the resolution of the stored data and the capture rate.

At its highest resolution, the DGx can capture a single screen at a resolution of 1280x1024, at a rate of 25 frames per second. However, to capture all of the CCS screens of a small Ops Room team (e.g. RT1, TrackSup, SWC, ORO) each at full resolution, the sampling rate would drop to 1.5 frames a second. In this quad mode all four screen could be replayed in separate windows on a single, high-resolution monitor. Intermediate resolutions allow for the capture of two full screens with playback on two separate monitors.

A brief evaluation of the DGx highest resolution data capture mode, that is, a single screen at 1280x1024 resolution, clearly showed that all of the video capture criteria could be met. In fact the captured image was visually indistinguishable from the original.

Since the quad capture condition represented the most probable mode that would be used to collect MOP data in small teams, it was decided to conduct the balance of the evaluation using this mode.

5.4.1 Set Up

The DGx comprises two units that can be vertically stacked on a desktop; one is the signal processor the other the data recorder (tape or digital). The DGx needs to be set up in close proximity to the workstations to which it will be connected in order to avoid long cable runs with the possibility of signal degradation.

Video connection: the DGx interfaces with the monitor input cable to each workstation and requires a dedicated cable to wherever the DGx is located. Since the DGx intercepts the video signal going to the workstation monitor, *a video splitter must be inserted into the line* in order for the video to be displayed on the NCOT workstation. Video splitters do not come as part of the DGx package and will need to be separately purchased; a video splitter will be required for each station monitored. The cost of a high quality video splitter is between $250-$450 US. High quality video splitters are designed for minimal signal degradation and would be a preferred option as opposed to a simple "Y" connection in the feeder line. In the assessment we evaluated the DGx recorded signal without a video splitter, which was not available.

Audio connection: The DGx has inputs for a left and right audio channel only, which need to be interfaced with the audio recorded at a workstation. In NCOT the audio is recorded in each position using a separate PC with an audio capture/digitising card. Attempts to interface the DGx directly with this card were not successful. An alternative would have been to use a cable splitter but none was made available by the DGx sales representative. In any event, with only two channels of recording available, audio capture in this way would probably be insufficient for T&E purposes. In order to capture simultaneous audio from all of the workstations and members of a small Ops Room team, a workaround would have to be constructed. The simplest way to do this would be to capture all spoken comms through an intercept of the microphone line at each workstation and to use an audio mixer to reduce these down to two tracks, which would then be input to the DGx for recording in synchrony with the graphics data.

The DGx unit itself needs to be connected to a separate (laptop) computer that runs specialised software to set up and record the data from the video sources of interest.

Assuming that all cables and accessories are available, it is estimated that the total set up time would be 30-45 minutes, and that this could be achieved by someone with minimal training. Once set up, the DGx system would make very little intrusion on standard NCOT protocols and configuration.

5.4.2 Evaluation

Video quality. This evaluation was conducted in quad mode with data captured from each of four workstations (RT1, TrackSup, SWC, ORO) for play back on a single monitor. Ultimately the data quality here is determined by the size and screen resolution of the playback monitor. Using an NCOT monitor, a*ll of the criteria outlined in 2 above were met* (as determined by simple visual observation and comparison with the original display), with the following two exceptions. First, track numbers in the radar area were not quite legible, although their associated symbology was clear. Second, diagonal lines took on a jagged appearance. Neither of these is considered to be a serious limitation for the collection of MOPs. In the case of the former, the data for any hooked track could be readily determined from the tabular data area and we do not see a special requirement to be able to read every track number. Further, the use of a larger, higher resolution monitor would probably allow the individual track numbers to be read.

Frame rate*:* as noted earlier, in high resolution/quad mode, DGx sampling drops to 1.5 frames a second. We believe that this will be adequate for most MOPs that are time based, since it is likely that the Navy is more likely to be interested in performance metrics with time scales that are much longer than fractions of second. Hence, the imprecision in temporal timing resulting from this sampling rate is likely to be of little practical consequence.

Audio capture: For reasons outlined above, no evaluation of the audio component was able to be successfully conducted.

Ability to set flags: In data capture mode, a DGx software function allows the user to readily set and name flags through a simple input table. The time of the flag is recorded along with the name. The functionality is intuitive and minimal training would be required.

Playback: In quad mode, all of the screens captured displayed the recorded information in synchrony and it was easy to distinguish events, local user settings and actions across the four workstations. In particular, all data in tabular fields could be readily determined as could the running clock. The playback could be paused without any noticeable degradation in displayed data and events could be tagged and named. The record could be easily fast forwarded and rewound, but not with the data visible on the screen. The recording could be quickly advanced or moved to previously stored flags with impunity. The only limitation on data playback was that once stopped, the record would default to the start. However, this could be readily overcome by the judicious use of flagged events as markers for rapidly advancing to the.desired location.

Once captured, the data record can be replayed independently of the site of capture, through a dedicated monitor. However, the original data record must be replayed through the DGx processor, which would necessitate relocating the unit to the site where data are to be analysed. The DGx allows for the record to be converted and output in an S-video format, which would allow the record to be stored and played back using standard video recording technology. However, such a format would probably not be suitable for data captured in quad mode because of the loss in critical detail when converting to the lower resolution video format.

Data management: The DGx can be purchased either with hard drive or tape storage, with capacities of 9 hours (per swappable hard drive) and 3 hours (per individual tape cassette) respectively. The only advantage of the tape storage format seems to be an ability to view the recorded data while in "fast forward" mode. Digital storage allows for easy data backup and management so that even when a disk gets full, its content could be saved to digital tape and then retrieved as required for review and analysis. The flexibility, capacity and management of the data storage system appear to meet all of the needs of T&E.

Portability: The DGx processor is approximately 18x17x3.5 inches and weighs 21 lbs and the digital storage unit is of similar dimensions and weight. Accordingly, the units would be readily transportable between test and analysis sites and are sufficiently robust to endure such handling. For installation purposes at the test site, only a small desktop surface is required.

Software functionality: The supplied software runs in a PC/Windows environment and uses fairly standard Windows interface style and conventions. The software appears to be intuitive and the basic functionality appears easy to learn. It is expected that there would be no requirement for additional specialised software support in order to use the system. The system configuration appears to be extremely flexible allowing for the interface with displays using a wide variety of hardware formats, resolutions and scan rates.

5.5 Conclusions, limitations and recommendations.

5.5.1 Conclusions

Based upon this assessment, we believe that the DGx system is a viable solution for the capture and analysis of high-resolution video images from NCOT workstations, when configured in individual or team (four player) format. The flexibility, data quality and the ability for quad playback in synchrony exceed existing NCOT functionality and meet the needs for T&E analysis. The portability of the equipment allows for off site analysis, which will require only a PC and a high resolution, large screen monitor. The control over the playback record meets the stringent requirements for the detailed and painstaking task of extracting MOP data from the video record.

The major limitation of the system is the requirement to add custom, ancillary equipment to multiplex the many audio sources that make up typical comms in an Ops Room into a single (or dual track) audio data record. In theory, this should be readily accomplished, but we are not in a position at present to know if the approach proposed above represents a viable solution to multiple audio source data capture and synchronised playback. (Note, that in any event there is a severe limitation with the existing NCOT system, whereby audio playback can only be done from one workstation at a time, thus precluding full team comms being replayed in co-ordination with the video).

5.5.2 Recommendations

If the MOP program to support COMDAT proceeds with data collection in NCOT, it is recommended that DRDC Toronto in conjunctions with Human*systems* Inc. explore ways to acquire the necessary DGx hardware for future use in the COMDAT MOP program. It should be noted that in addition to the basic DGx system video splitters will also have to be purchased.

Notwithstanding the utility of the DGx system for data capture in the NCOT environment, there are two other viable applications for the system that could support the COMDAT program. First, for data collected for MOP purposes at the ORTT (see reference 2), the DGx system could be readily interfaced with the ORTT playback unit to allow MOP data to be taken off site (e.g. to DRDC-Toronto) for more detailed analysis.[20] This would save on local accommodation costs for the T&E analysis team. Second, for COMDAT trials planned for the CSTC, the DGx could be readily interfaced to capture data in a manner suitable for MOP analysis that is not feasible with the

[20] Assuming security requirements could be fully met.

existing CSTC data capture technologies. Such a capability would allow a more direct means of before and after comparisons of the potential impact of COMDAT MSDF technology.

5.5.3 Other applications

Finally, it should be noted that the DGx system is capable of capturing data from a wide range of military C2 environments and hence could be useful to DRDC-Toronto in any future work involving radar displays, sonar displays or teams communicating and interacting through networked workstations. These capabilities may also be of interest to other DRDC partners, who may be willing to participate in the purchase cost.

The system should also be of interest to the *Navy* in providing a major enhancement to its current capabilities to capture data for detailed analysis of system performance, whether for C2 purposes or to evaluate any other interactive workstation context. The DGx would allow for a transparent, non-obtrusive implementation in a variety of real-time, data capture and analysis environments, ranging from the ship to land based test facilities.

5.6 Costs

The cost of the DGx system components is as follows:

Part Number: RGDGx100: Description: DGx Recorder and Replay unit
OEM List: $32,353 CDN

Part Number: RGDTR20P:Description: 3 Hour Tape Deck
OEM List: $8,205 CDN

Part Number: RGDSS120GB: Description: 120 Gb Hard Disk storage (replaces above unit)
OEM List: $8,205 CDN

Part Number: RG7207148 :Description: Virtual Control Panel Software
OEM List: $798 CDN

The total cost of a full system with the hard drive storage would therefore be: $41356 to which would need to be added the cost of four high bandwidth video splitters at an approximate cost of $2000.

5.7 Alternate approaches

An alternate approach would be to use a video scan converter and high quality, video data storage device at each separate workstation. Such a system would cost approximately $10,000. Since such a unit could only record from a single workstation at a time, the costs for a four unit recording capability would be almost the same as for the DGx system. Further, such a system would have two major disadvantages compared with the DGx, namely lower resolution and lack of synchronised playback. Note that to play back in synchrony would require the user to simultaneously hit four "play" or "stop" buttons - thus with progression though the data record, it is likely that the tape synchronisation would increasingly suffer. This hardware option would however have some viability if there were interest in recording data from one or possibly two workstations, and there may be some possibility of gaining access to units already purchased by the Navy.

5.8 References

1. M. L. Matthews, R. D. G. Webb, A. R. Keeble. *Findings of the Test and Evaluation Proof of Concept Trial at the NCOT Facility*: April 2002. Section four, this report.

2. M. L. Matthews, A. R. Keeble., L. Bruyn. *Follow- up evaluation of the Operations Room Team Trainer as a Test Environment of collecting MOP data to support COMDAT*. September 2002. In: Matthews, M.L., Keeble, A.R., Bruyn, L.E., Webb, R.D.G. and Lamoureux, T.M. (2003). Development of measures of performance for evaluating the COMDAT technology demonstrator: potential use of training records from the Operations Room Team Trainer (ORTT). DRDC Report No. CR-2004-010.

6. Technical Memorandum:

DISCUSSIONS WITH MDA CONCERNING NCOT FUNCTIONALITY TO SUPPORT TEST AND EVALUATION ASSOCIATED WITH THE COMDAT PROGRAM

6.1 Summary

6.1.1 Background

This Technical Memorandum was initiated to investigate further the deficiencies identified during the POC. The Technical Memorandum contained:

- A list of required enhancements to support T&E;
- Summary of initial discussions between HSI® and MDA;
- Subsequent MDA comments;
- Further MDA comments;
- Suggested next steps.

6.1.2 Findings

Findings are grouped under four headings: scenario creation; scenario execution; scenario playback and analysis; and miscellaneous.

6.1.2.1 Scenario Creation

There were a total of three enhancements listed under this section. MDA stated that it would address issues identified by the Department of National Defence (DND) as time and priorities permitted. HSI® felt that it would be useful to understand the process by which this was achieved. Subsequently, problems identified with scenario creation appear to have been corrected but this would require verification before any data collection were to be undertaken.

6.1.2.2 Scenario Execution

There were a total of two enhancements listed under this section. The ability to add a time-stamped flag was not thought by MDA to be difficult to implement. However, data capture at the keystroke level (the 'Team Assessment' feature) was potentially problematic due to performance issues. HSI® was to investigate the Team Assessment feature at a future meeting with MDA. To assist in this process, HSI® undertook to compile a list of desirable keystroke level recordings. HSI® felt that performance problems could be minimised if data capture took place when no other software was being run. MDA has, at the time of this Technical Memorandum, not yet commented on how data was stored and made available for analysis.

6.1.2.3 Scenario Playback and Analysis

There were a total of three enhancements listed under this section. Importantly, there may be security issues associated with taking data away for analysis. MDA state that synchronised

playback at individual workstations is possible. However, there is no control over playback except to start and stop playback. A Commercial-Off-The-Shelf (COTS) tool that allows fast forwarding, rewind, etc. can be implemented to work with NCOT, but is not available to typical NCOT users and can only work on a single workstation. To extend this to multiple synchronised wokstations would represent a significant effort. MDA has no plans to do this.

The issue surrounding control over playback to facilitate analysis could be addressed by the use of the DGx recorder. Even with this solution, there would still, however, be no ability to capture real-time key presses with a time stamp due to the system performance costs associated with implementing such a function.

6.1.2.4 Miscellaneous

There were a total of eight enhancements listed under this section. They focused on a number of small enhancements to assist novice users to distinguish different features of the display and to use audio recordings.

6.1.3 Conclusions

Original concerns with the ability to pre-program and automate scenario events appear to have been addressed. A major outstanding issue concerned the ability of NCOT to provide control and manipulation of the scenario replay to facilitate data analysis. The original requirements still appeared to be difficult to achieve with the current NCOT software. To obtain the level of control over playback data to allow the efficient extraction of T&E data would require significant new software functionality, or the use of supplementary data capture systems, such as the DGx Spectrum.

A second major outstanding issue concerned the lack of a capability to capture real-time, specific key presses at the ORO and other team workstations, which are time stamped during scenario execution and recorded to a data file. Functionality to capture such events would require MDA to develop additional software and there were concerns that if the list of events to be captured were lengthy, there could be some impact upon scenario execution performance.

6.2 Introduction

This Technical Memorandum addresses the results of discussions with MDA concerning the need for NCOT software and hardware enhancements to support T&E for COMDAT MOP. It is based upon the evaluation of the NCOT capabilities reported in reference 1 and a subsequent meeting between a Human*systems* Inc. consultant and an MDA software specialist in October 2002. Initially, it had been intended for this to be a working meeting involving a hands-on evaluation of the software elements under discussion, using scenario vignettes. However, with a recent shift in the project towards conducting future T&E trials in the ORTT, this meeting was largely intended to obtain verbal clarification from MDA concerning some of the major software issues.

The following contains four components:

1. A list of the required enhancements to software functionality to support T&E, organised into related themes (taken from reference 1).

2. HSI® notes: A summary of the initial discussion held between Michael Matthews and Troy Yee of MDA concerning these requirements.
3. MDA comments: subsequent MDA comments and notes provided by Troy Yee
4. MDA supplementary comments arising out of the meeting in October.
5. Suggestions for next steps

6.3 List of T&E software enhancements

Scenario creation

1. *Provide a capability in the master scenario file to allow entities to follow a pre-scripted trajectory. This may be readily accomplished by fixing the problems in implementing the NCOT Scenario Builder function of "**Pre-determined tracks**". (NCOT Manual Section 6-27).*
2. *Allow entities to automatically enter into the scenario at pre-specified times and follow a constant track and speed.*
3. *Allow entities to automatically enter into the scenario at pre-specified times and follow a pre-scripted trajectory.*

HSI Notes

The existing functionality should support items 1-3. However, if it is unstable there is a workaround. The "save as" instruction should allow a scenario which has been created and manually implemented at the instructor station to be saved. It can then be subsequently run and the entities will automatically enter the scenario and behave in the manner in which they had been manually programmed. Note behaviours will be limited by the inherent dynamics of the entities - i.e. they cannot do things that are not within their normal functional capabilities.

MDA Comments

MDA will address scenario builder software deficiencies that prevent successful use of currently provided capabilities, particularly those that lead to unexpected program termination and subsequent loss of data. Any such issues identified to MDA by the DND will be addressed as time and priorities permit.

NB. When using the 'Save As…' feature, ownship manoeuvres will also be saved and replayed. If your experiments include subject-induced changes to ownship motion, it would be best to save a scenario that does not include ownship manoeuvres.

MDA Follow-up comments

MDA indicated that there were in fact some bugs in the existing build, specifically related to some aircraft altitude changes, but that other than that this "pre-determined track" function works. The existing bugs should be cleared up with the most recent build, *so this issue should now be resolved.* MDA also noted it should be possible to have an aircraft appear on a particular course and at a specified altitude immediately, as opposed to having it "take off" and climb to a specified altitude. A further capability exists to "build" hills etc into the game's geography, (which would mask the radar return of aircraft flying on the other side).

Next Steps/Comments:

These software problems now appear to be corrected, however they should be verified practically before any data collection trials are begun.

Scenario execution

4. *Provide a capability for a time-stamped flag to be entered from an instructor workstation into the scenario record. This could be achieved by capturing function key presses and logging the time of the press and saving to a file. The resulting data file should comprise a time-ordered list of the key presses and time (scenario time) when pressed. This file should be in an ASCII text/tab delimited format to allow for export to a spreadsheet.*

HSI® notes: none

MDA Comments

This functionality is not currently supported but would not be difficult to add if requested. See also next response.

5. *Provide a capability to capture specific key presses at the ORO and other team workstations. These are to be time stamped and recorded to a data file in the manner described under item 4. Specific actions required to be captured include: quick action buttons (QABs), numeric keys, "enter" key, F2 and the alarm button.*

HSI Notes

The next software version (September release) will support team assessment whereby an entire event will be logged and timed (e.g. missile launch-detect-response). These will be specifically coded and not amenable to any further analysis. Verbal comms will not be recorded as part of this.

In general, the system has low bandwidth for further data capture and logging. Therefore we should identify which specific keyboard and QAB events would be high priority for capture. It would not be feasible to capture a large range of these (e.g. every keystroke).

MDA Comments

The team assessment feature captures the occurrence of specific events. Upon completion of a game, the events are saved to a log file at the Instructors discretion. The results of several specific assessment calculations are also included in the log file. The event data (game time of event with application supplied description/data) in the file is available for further analysis once saved. Additional events can be added but require specific application coding (instructor initiated events are possible).

The team assessment feature is designed to capture events that occur relatively infrequently. Capturing and logging events that occur too frequently could adversely impact a running game and potentially result in very large data files. File I/O during a running game is avoided due to potential performance impacts. These issues are of particular concern in situations where multiple concurrent games are running within a facility.

MDA Follow-up comments

The team assessment feature captures only macro events, such as the times that an entity was activated, when a track was generated, the target designated to STIR, missile launch etc. It does not capture individual operator or team actions that precipitated these macro events.

Next Steps/Comments:

If NCOT is to be used for future data capture trials, HSI will generate a basic list of keystroke/QAB etc actions to support T&E, based on the detect to resolve process.

Concerns over adverse effects on system performance may be ameliorated if T&E is conducted at times when no other software is being run in NCOT for training or other purposes.

MDA has still to comment on how the captured data may be stored and made available for analysis.

Scenario playback and analysis

6. *Develop a capability for control over playback files to include: slow motion, pause/re-start, rewind, fast rewind/fast forward and go to flags (as in#4).*
7. *Provide an ability to download data files to external media.*

8. *Provide an ability to replay multiple files simultaneously on different workstations (preferably in synchrony)*

HSI Notes

Re: item 7: This can be done - however, we would need to check with Navy re security issues. The resulting file could be played back on a local HP workstation but this would require the purchase of a COTS tool to support this.

There is no capability at present to convert the proprietary video recording format into standard formats (e.g. MPEG). The COTS tool claims to be capable of converting stored video data to MPEG format but initial attempts by MDA to perform such a conversion were unsuccessful. MDA had originally sought a hardware solution for this but Navy preferred a software approach.

The existing playback system is a COTS product that does come with a playback tool that has a less than optimum user interface and may be somewhat "buggy". The tool is not currently supported and has no special NCOT interface. The tool allows for fast forward, slow motion pause, restart and rewind of video. It can also go to specific points and loop. The tool works on an individual workstation. The tool does not support audio playback, which would have to be handled manually.

In addition, the instructor "Debrief" tool allows for the co-ordinated playback of scenario record files from up to four workstations to be played back on a subset of NCOT workstations. Maintaining synchrony through playback would require manual control and co-ordination by operators at each workstation using the replay tool local to that workstation. Software could be written to accomplish this automatically, however, it is not in the current plan.

MDA Comments

NCOT currently supports 2 methods of playback for audio and video recordings. The ‘least effort’ method allows for ‘in situ’ playback – audio and video recordings made on a specific workstation can be played back on that workstation. MDA has provided an interface for this option on the instructor station (“Playback” option of the “Desktop” menu of the Classroom configuration application). If recordings were made of several stations participating in a single game, these recordings can be played back simultaneously at each station in near synchrony. In this mode, the recordings will play from beginning to end with no other control available except to stop playback.

The second method provided involves a two-stage process. The operator must select recording sessions to store to a ‘debrief catalogue’ using an MDA provided interface on the instructor station (“Save files for debrief” in “Record” option of “Desktop” menu of the Classroom configuration application). After the files are saved, the operator may play these files back at any other station using the MDA supplied ‘Debrief’ application. This application allows for playback of a single recording session at the station where the application is running. Again, no control over the playback is available other than to stop playback. After using the ‘Debrief’ tool at a station, the files associated with the recorded session are available for playback using the ‘Playback’ option (first method above).

The COTS tool that allows playback of video at arbitrary speeds from/to arbitrary points in the recording is not available to typical NCOT users. This tool only works on a single workstation. It might be possible to create an interface that permits co-ordination of instances of this software running on multiple stations to support concurrent playback with pause/speed/loop capability. This would not be a trivial task.

Audio can be played back on any platform that supports .WAV format.

MDA Follow-up Comments

Although, a second GUI exists within the COTS tool that has some playback features, MDA cannot at this time provide any details of its functionality.

Next Steps/Comments:

The MDA response suggests that there are still a number of remaining deficient critical requirements to support the replay and analysis of T&E data. To provide the level of support required would require some degree of programming effort and MDA has no current priorities for this.

This reinforces the need to pursue in parallel other avenues for data capture that hold better promise for supporting the T&E analysis requirements.

Since the initial evaluation, NCOT has become a secure facility, hence any future T&E data sets showing performance implications may need to be classified in some way.

Miscellaneous

9. *Provide data file management capabilities*

HSI Notes

MDA states that this capability exists and could be provided by local network manager.

10. *Provide an override of the workstation "sleep" function or a much longer time before activation.*

HSI Notes

This can be done by the NCOT system administrator.

11. *On SSD enable dimming of symbology so that can operator can examine radar paint alone (as occurs on the actual CCS)*

HSI Notes

This can be achieved by turning off symbology through existing functionality - but not dimmed. This would be a satisfactory solution to a lower priority requirement.

12. *Provide a capability for the creation of CCS overlays ahead of time and to allow them to be integrated into the Tactical Situation Area of the SSD when the scenario is run.*
13. *Enable instructor to draw "lines" and save for later use e.g. patrol areas (e.g. like Word Draw options plus text insertions)*

HSI Notes

This is not supported now, but QABs allow save of CCS overlays. This capability could be possibly be enhanced. HSI will check out this capability at such time when future data collection in NCOT is required.

14. *Provide a means of distinguishing between active tracks and stationary or parked tracks in the instructor screen i.e. to rapidly identify all tracks in play at the moment*

HSI Notes

Instructor workstation provides overlays of velocity vectors and history trails. This should help to distinguish active from stationary tracks. This issue appears to be resolved, since items 1-3 from the requirements list appear to be addressed appropriately.

15. *Allow audio playback on any workstation*

HSI Notes

This is provided through Debrief tool, which allows playback on any NCOT workstation and not restricted to the one on which audio was originally recorded. It is also possible to play the .WAV file associated with an audio recording on any .WAV capable platform. Mechanisms exist to get access to these files.

16. *Improve signal to noise ratio (sound quality) of audio playback.*

HSI Notes

This is a problem within the COTS tool itself, which masters all audio data onto a single track for playback. There is no simple solution. Prior to any future data collections in NCOT, HSI will check out adjustments to audio gain and ways of minimising intrusive noise from alarms.

6.4 Conclusions

These conclusions are based upon the above discussions with MDA and the information provided therein. Should the T&E program require future data collection in NCOT, some of the conclusions will need to be further validated in context using scenario vignettes.

- Original concerns with the ability to pre-program and automate scenario events appear to have been addressed.
- One major outstanding issue concerns the ability of NCOT to provide control and manipulation of the scenario replay to facilitate data analysis. The original requirements that we have identified still appear to be difficult to achieve with the current NCOT software. To achieve the level of control over playback data to allow the efficient extraction of T&E data will require significant new software functionality, or the *use of supplementary data capture systems*, such as the DGx Spectrum (see reference 2).
- A second major outstanding issue concerns the lack of a capability to capture real-time, specific key presses at the ORO and other team workstations, which are time stamped during scenario execution and recorded to a data file. Functionality to capture such events would require MDA to develop additional software and there are concerns that if the list of events to be captured were lengthy, there could be some impact upon scenario execution performance.

6.5 Next Steps

Given that the project is now looking at the ORTT as the primary source of data capture for future T&E trials, it is recommended that any future options for either NCOT software revisions, or the purchase of a supplementary data recording capability, be put on hold until such a time as the issues become relevant to the program.

6.6 References

1. MATTHEWS, M. L., WEBB, R.D.G. and KEEBLE, R. Human*systems* Incorporated Technical Memorandum: *Findings of the Test and Evaluation Proof of Concept Trial at the NCOT Facility*: April 2002. Section three, this report.

2. MATTHEWS, M. L. and KEEBLE, R. Human*systems* Incorporated Technical Memorandum: *Evaluation of the RGB Spectrum DGx Digital Graphics Recording System as a means of collecting NCOT MOP data to support COMDAT.* August 2002. Section six, this report.

THIS PAGE INTENTIONALLY LEFT BLANK

7. Overall Summary

The work reported above provides a comprehensive package of MOPs for assessing the impact of the COMDAT TD on Ops room teams in the HFX class environment, and was predicated initially on conducting T&E trials in NCOT. In a subsequent POC in NCOT intended to assess the practicality of gathering MOPs, several limitations for T&E purposes were identified in the NCOT functionality. A number of corresponding suggestions to improve the gathering of MOPs were proposed and discussed with the contractor in charge of NCOT support (MDA). Also, a candidate add-on recording system was assessed for its ability to assist in the collection and more efficient analysis of MOPs.

During the course of this work it became increasingly apparent that the ORTT might offer a better environment for COMDAT T&E purposes, for a number of reasons. First, a number of complex scenarios had already been developed in the ORTT for training intact Ops Room teams. Second, the ORTT tasks associated with the detect-to-resolve cycle are performed in the appropriate context of a fully functional Ops Room, thereby representing all of the normal C2 interactions that take place under operational circumstances. (Unlike in NCOT, where the isolation of a small sub-team from the overall operational environment not only reduces workload but also excludes many relevant events and communications that impact even upon the simplest of tasks.) Third, the data recording system provides co-ordinated capture of all Ops Room CCS screens and communications and allows for synchronised playback. Fourth, the possibility of extracting MOPs from existing data records seemed a promising option, given the logistical difficulties and complexities of mounting a T&E trial dedicated to just COMDAT MOP data collection. Fifth, the data obtained from teams undergoing training in the ORTT was seen to have greater reliability and validity than anything that might be obtained from a dedicated T&E trial in NCOT.

Therefore, it was recommended that the continuing development of MOPs predicated on the use of NCOT be suspended until a later date pending a more formal assessment of the ability of ORTT training records to support the extraction of data relevant to COMDAT MOPs.

These technical reports and memoranda associated with the follow-up work in the ORTT are reported in a separate, but parallel, Technical Memorandum/Report entitled: Development of Measures of Performance and a Trial Plan for Evaluating the COMDAT Technology Demonstrator: Potential Use of Training Records from the Operations Room Team Trainer (ORTT) for data collection .

THIS PAGE INTENTIONALLY LEFT BLANK

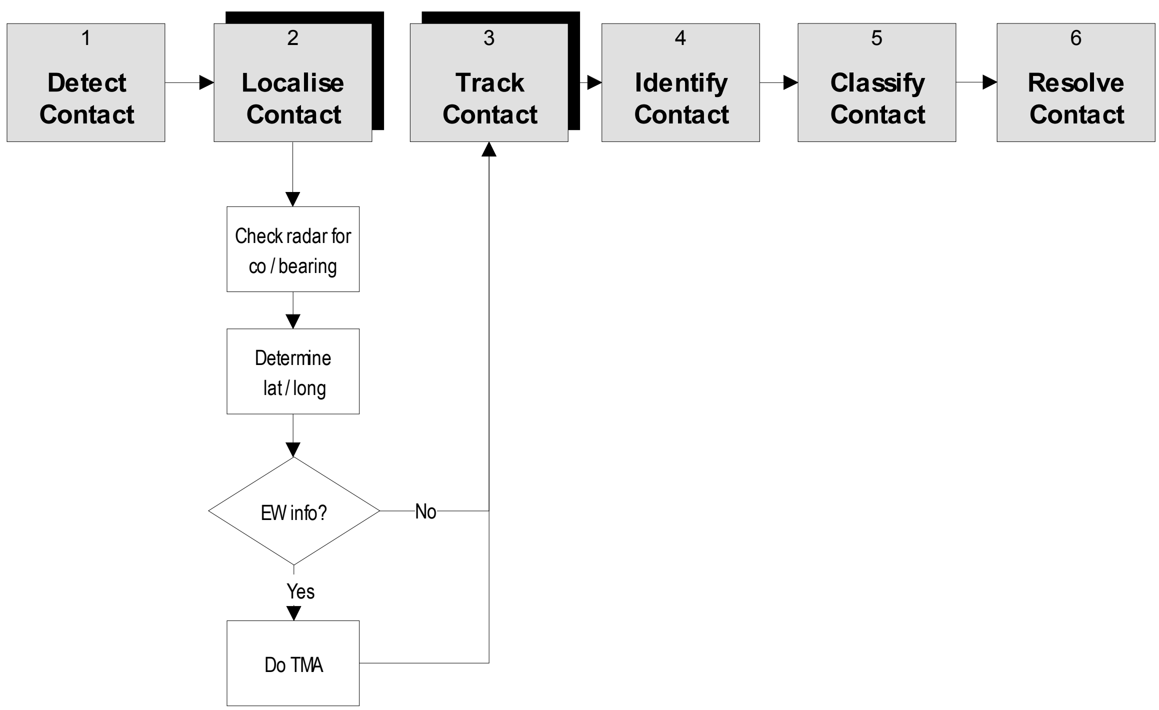
8. List of Acronyms

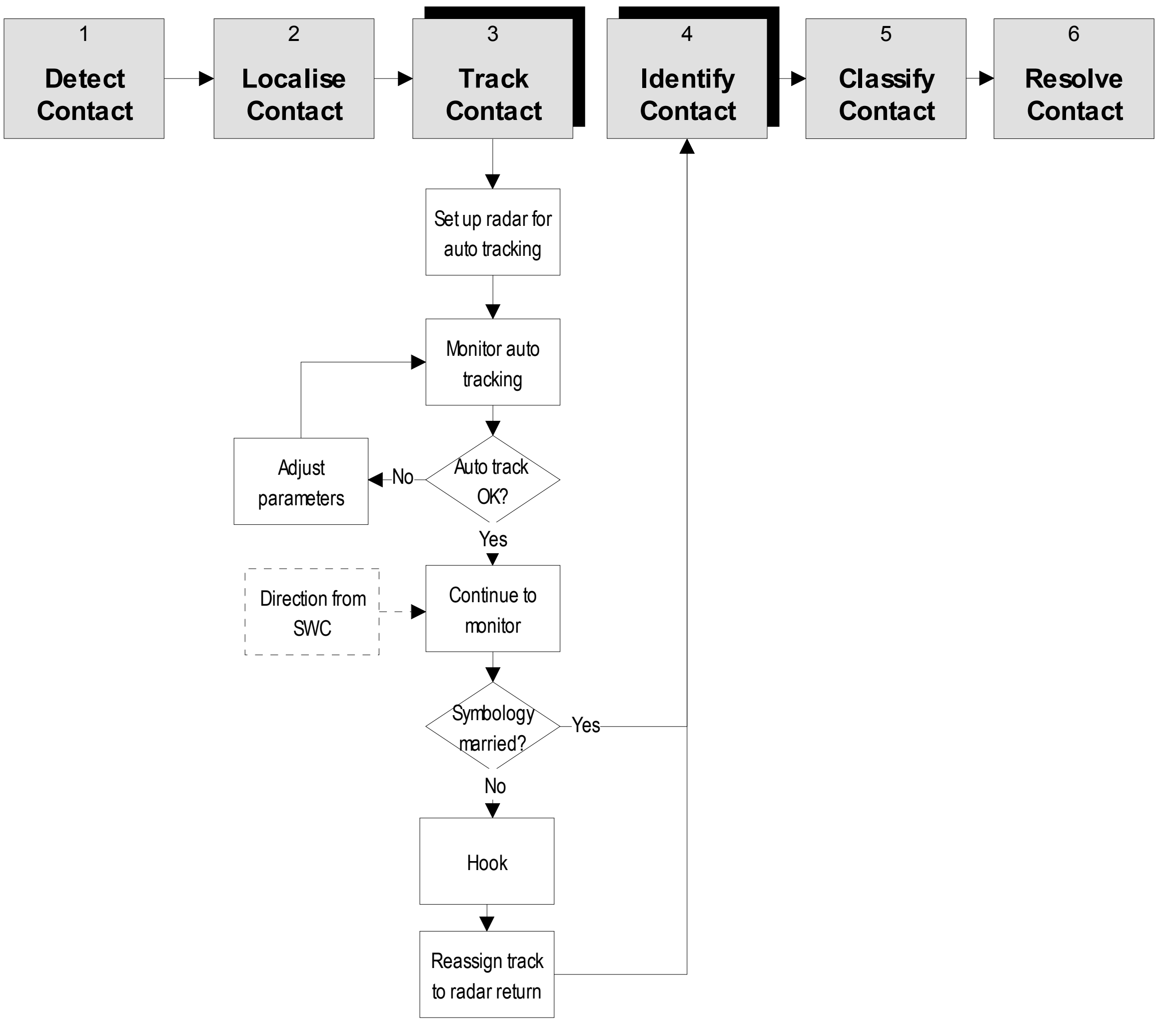
AAW	Anti-Air Warfare
AWW	Above-Water Warfare
A/C	Aircraft
AOI	Area Of Interest
SWC	Sensor Weapons Controller (also Surface Warfare Commander)
CO	Commanding Officer
COP	Common Operational Picture
CPA	Closest Point of Approach
GCCS	Global Command and Control System
LAP	Local Area Picture
MCOIN	Military Command Operational Information System
MOP(s)	Measure(s) of Performance
MSDF	Multi-Sensor Data Fusion
MSP*	Maritime Surface Picture
MSubP*	Maritime Sub-surface Picture
MTP	Maritime Tactical Picture
NCOT	Naval Combat Operator Trainer
OPGEN	General instructions from the Tactical Commanding Officer to Commanders
OPTASK	Operational instructions from the appropriate commander that detail the conduct of operations (note: there will be a number of different OPTASKs for different warfare areas, such as air, surface and subsurface).
ORO	Operations Room Officer
ORTT	Operations Room Team Trainer
PU	Participating Unit
RAP	Recognised Air Picture
RLP	Recognised Land Picture
RMP	Recognised Maritime Picture
SITREP	Situation Report
SWC	Sensor Weapons Controller (also Surface Warfare Commander)
TDP	Technology Demonstrator Project
T&E	Test and Evaluation
TG	Task Group
WAP	Wide Area Picture
WD	Warfare Director

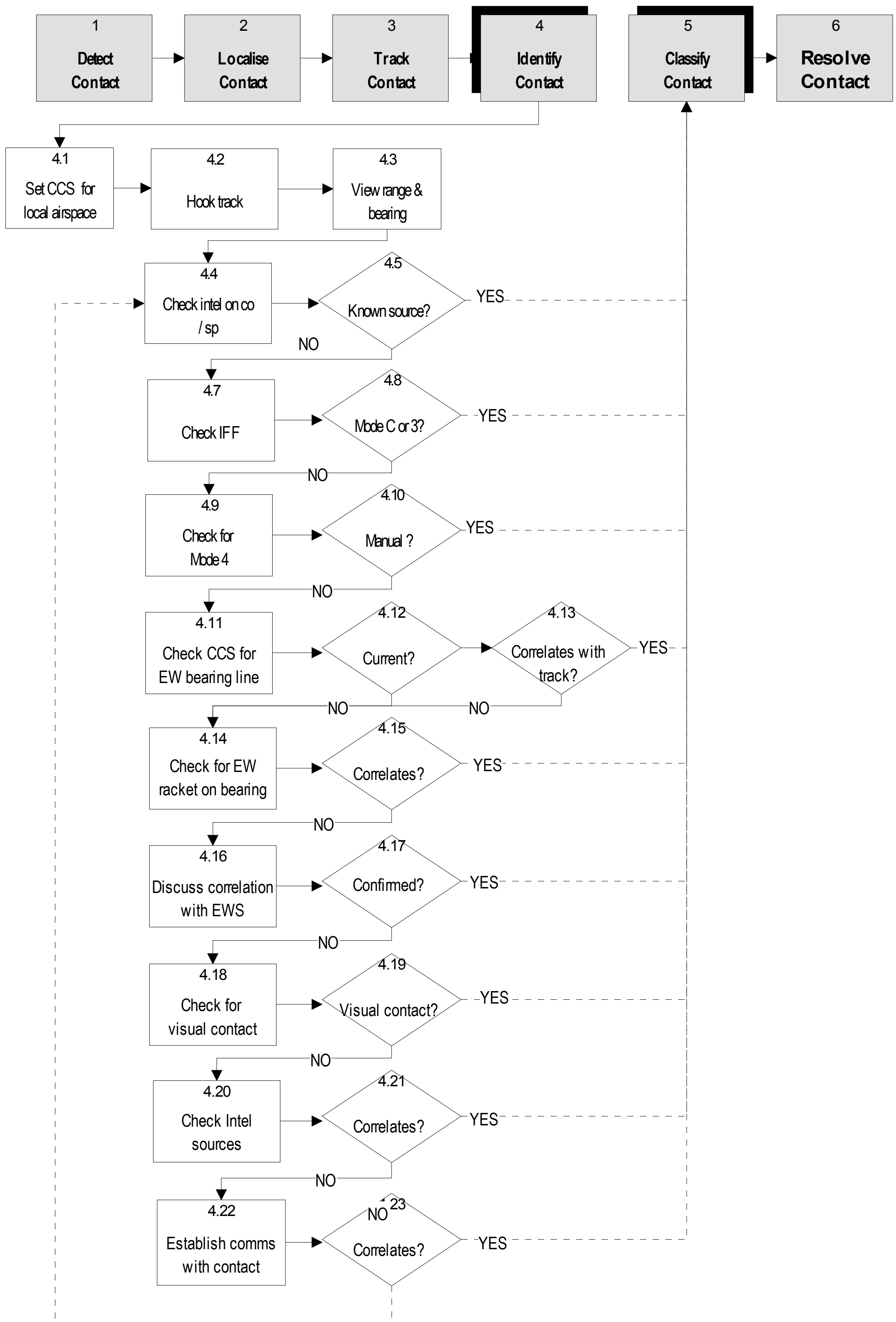
*Note: these are abbreviations coined for present purposes and may not represent current Navy acronyms.

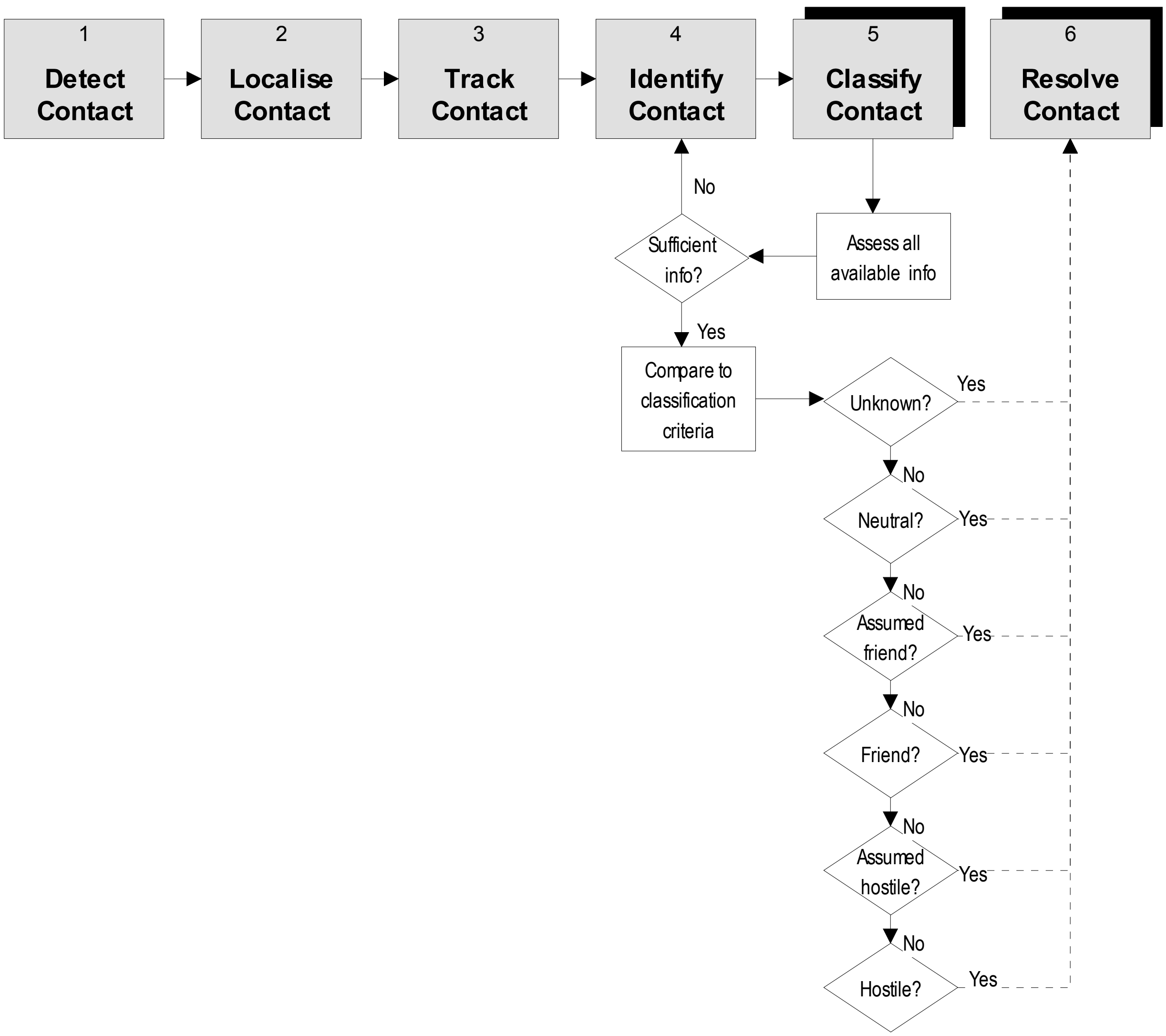
THIS PAGE INTENTIONALLY LEFT BLANK

Annex A:
Information Flow Diagrams of the Detect-to-Resolve-Cycle




1
Detect Contact
2
Localise Contact
3
Track Contact
4
Identify Contact
5
Classify Contact
6
Resolve Contact
Set up radar for auto tracking
Monitor auto tracking
Auto track OK?
No
Adjust parameters
Yes
Direction from SWC
Continue to monitor
Symbology married?
Yes
No
Hook
Reassign track to radar return


1 Detect Contact
2 Localise Contact
3 Track Contact
4 Identify Contact
5 Classify Contact
6 Resolve Contact
4.1 Set CCS for local airspace
4.2 Hook track
4.3 View range & bearing
4.4 Check intel on co / sp
4.5 Known source?
YES
NO
4.7 Check IFF
4.8 Mode C or 3?
YES
NO
4.9 Check for Mode 4
4.10 Manual ?
YES
NO
4.11 Check CCS for EW bearing line
4.12 Current?
4.13 Correlates with track?
YES
NO
NO
4.14 Check for EW racket on bearing
4.15 Correlates?
YES
NO
4.16 Discuss correlation with EWS
4.17 Confirmed?
YES
NO
4.18 Check for visual contact
4.19 Visual contact?
YES
NO
4.20 Check Intel sources
4.21 Correlates?
YES
NO
4.22 Establish comms with contact
NO
23
Correlates?
YES


1
Detect Contact
2
Localise Contact
3
Track Contact
4
Identify Contact
5
Classify Contact
6
Resolve Contact
No
Sufficient info?
Assess all available info
Yes
Compare to classification criteria
Unknown?
Yes
No
Neutral?
Yes
No
Assumed friend?
Yes
No
Friend?
Yes
No
Assumed hostile?
Yes
No
Hostile?
Yes

Annex B:
Follow-up Evaluation of the NCOT Facility

Notes from the HSI® visit to NCOT November 21/22, 2001 concerning MOPS and Test and Evaluation needs for COMDAT1 Evaluation.

1. Sufficient workstations can be made available to meet T&E Needs.
2. Workstations can be configured and networked to simulate a team comprising the RT1, SWC and ORO.
3. Workstations are available and can be configured for the T&E team to monitor the workstations of the Ops Room team.
4. Workstations are available and can be configured for the T&E team to manipulate scenario events.
5. Workstations screens and all communications can be captured and stored on a hard disk for later replay. The capacity limitation has not been tested to date. Audio data are captured to a PC file, video data to the HP workstation and is limited by hard disk size. Hard disks cannot be swapped during a session to enhance capacity. The current video limit is thought to be about one hour with the existing size of disks. This needs to be more rigorously tested with a view to determining the actual limit, what trade-offs in captured functions may have to be made and whether a larger disk will solve the problems for T&E data collection.
6. Flags for T&E purposes are also on DND's wish list but MDA has no current plans for implementing these. In fact, the COTS proprietary supplier is not interested in doing it. As to timings, this might be possible to refine the software because MDA controls some aspects of that. The best approach would be to come up with a short wish list of timing start and stop points for T&E purposes and discuss informally with MDA, who could probably program a menu for each, provided they can be uniquely defined.
7. Audio / video playback: the record system uses proprietary COTS, thus information cannot be played back either on any other machine (e.g. off site at DRDC or HSI®) unless a licensed system is available with installed software.
8. Need to explore with DRDC/MDA the cost and logistics of acquiring a system that will run NCOT software and the requirements for installation and maintenance.
9. Radar tracks cannot be simulated on the CCS, hence there is a need for an actor to play the role of the RT1 to perform the task of creating tracks. Similarly, an actor will need to be present to perform the role of the SWC.
10. There are no existing scenarios that can be suitably adapted for T&E purposes. However, the library of already created game entities (e.g. aircraft, ships and their associated attributes and kinematics) can be used for scenario building and hence these will not need to be created from scratch. A brief review of these entities suggests that they are sufficiently diverse in type and number to satisfy the scenario building requirements for T&E.

11. Existing land mass geographical maps are suitable for the purposes of T&E. However, they represent a flat two dimensional surface only, hence desirable elements of the operating environment such as coastal hills and valleys cannot be simulated in terms of their effect on the radar picture. This will mean that the loss of radar signal and signal degradation that would normally occur from air contacts moving in such a region will have to be simulated in real time by a member of the T&E team.

12. In the current software build, the behaviour during the scenario of pre-programmed events is unreliable. The next build scheduled for delivery in January is supposed to rectify this. Currently, there is a lack of trust that the software will indeed handle pre-programmed events in the required manner. Therefore it has been recommended that real time control by the T&E team of some events can be anticipated. This will create an additional administrative overhead in running the scenarios and will require the development of a careful, detailed script that is well rehearsed with highly trained personnel.

13. T&E personnel are able to control the behaviour of entities in real time by making them active or inactive and moving them to new geographical positions while inactive.

14. The facility appears capable of being able to reproduce all of the kinds of events anticipated in the initial scenario plan. Some can be readily implemented, others will require some workarounds.

15. After developing the master events list and scenario schedule, the implementation will require some iteration in testing and refinement. This can only be conducted on an NCOT workstation. The skills to conduct this can be readily acquired by the T&E team.

16. The scenario building tools available are not very user-friendly and must be compensated for by spending more time in paper planning beforehand. The inaccuracy of the system in driving entities can also be compensated for by spending more time during the programming stage - i.e. playing it over and over until the entities do what you want. On the positive side, after developing a good background scenario, the work required to complete a whole scenario by inputting unique foreground events is minimised.

17. The facility is adaptable to adding an ancillary audio and video recording capability for T&E purposes.

Annex C: Scenario Description

General Situation

As a result of Yemen's continued refusal to take decisive action against Al Qaeda cells operating within its borders, President Bush has included Yemen as an element in the Axis of Evil. The USS George Washington Battle Group has been positioned in the Gulf of Aden and carrier based aircraft have been conducting overt operations to demonstrate air superiority in the Gulf up to the territorial air space of Yemen. President Bush has been working to build international support for military action against Yemen but has met with resistance, particularly from the Arab nations, Russia and China.

While coalition-building efforts continue naval forces from the United States, Canada, the United Kingdom and Australia have begun to monitor shipping in the Gulf of Aden and the Red Sea in anticipation of a UN resolution establishing a Maritime Interdiction Operation in the region. The main focus of the monitoring effort is in the Gulf of Aden as concerns mount that ships from Iran may be transporting Al Qaeda supplies and personnel to Yemen ports. HALIFAX is the only ship in the southern portion of the Red Sea. HALIFAX's role for the time being is to become familiar with the maritime environment in the area. If a MIO is authorized by the UN, HALIFAX will participate.

Yemen has claimed that its forces have been preparing for strikes against the Al Qaeda cells but there appears to be little resolve for action. Al Qaeda members are believed to be relatively free to travel within Yemen, and are known to cross borders into southern Saudi Arabia, Eritrea and Ethiopia. Refusing to be intimidated, Yemen military aircraft have on occasion intercepted US Navy aircraft in international airspace and have approached the George Washington to within 20 miles. They have also routinely approached HAL to within 5 miles. To date the Yemen aircraft have carried only air-to-air missiles; they have not carried bombs or anti-ship missiles. Yemen aircraft include Fitters (FBA), Fishbeds (intercept) and Mirage (anti-ship). The Fitters carry two 500lb iron bombs and the Mirage carry two Exocet ASMs. Yemen Tarantual patrol boats have routinely operated within territorial waters in the Gulf of Aden and Red Sea. The Tarantuals carry SSN-2C anti-ship missiles. There are two operational coastal missile batteries on the Red Sea coast and their associated Puff Ball radars are routinely active for extended periods of time. Tension between Yemen and the US has been relatively high since the attack on the USS COLE in Aden, with the September 11th attacks serving to heighten emotions even more. Although the Yemen government has made no threats, Al Qaeda elements within Yemen have made unspecified threats against American interests in the area.

To further complicate the situation, Yemen and Eritrea have begun a war of words over oil rights in the southern Red Sea. American oil companies based in Eritrea operate four oil rigs in international water just off of the Yemen island of Kubra. Yemen claims that the oil rigs are within their territorial waters but the international community recognizes the area as international waters. Yemen maintains a naval presence in the vicinity of the oil rigs, usually consisting of one or two Tarantual patrol boats. Yemen aircraft on occasion approach Eriteran airspace to exert pressure and probe defences.

HALIFAX is operating near the major shipping lane between the Suez Canal and the Gulf region. Large ships including oil tankers regularly pass through the Bab el Mandeb, the strait between Yemen and Eritrea/Djibouti. Other maritime traffic in the area consists of some small wooden fishermen and pleasure craft, supply boats going to and from the oil rigs from Aseb, Eritrea, and small, fast cigarette boat-like smuggler traffic between Yemen and Eritrea. These smugglers trade primarily in cigarettes and liquor, but intelligence indicated that they may also be transporting Al Qaeda personnel. The waters in the area are very polluted, with floating animal carcasses and other debris being commonplace.

Apart from the routine commercial air traffic in the region, helicopters fly between Aseb, Eritrea and the oil rigs, and numerous small private planes transit between Eritrea, Yemen and Saudi Arabia. Many of these planes are flown by rich Saudi's , the same demographic that sails recreational boats in the area.

HMCS HALIFAX is carrying one Sea King helicopter and has a full operational weapons load. The ship is manned at Remar and has no major operational defects. The ship completed a full RSP, including Workups, three weeks prior to deploying from Halifax. The ship has been on patrol for eight days and expects to remain another 12 prior to being relieved. HALIFAX will continue to rotate into the patrol area for the foreseeable future. HALIFAX is operating under national OPCON and is using Canadian OPTASKS and ROE. ROE include the authority to warn off aircraft up to and including Warning 5. HALIFAX is in the Second Degree of Readiness, with CCS is in Semi-Auto, STIRS in Hot Standby, and all missiles tuned.

T time is 1230 local. Port watch is just coming on watch for the afternoon. Skies are clear and visibility is good. The captain is on the bridge. The XO is in her cabin.

Rules of Engagement:

Self defence only. Use of FC radars for height finding is authorized.

Identification Criteria:

AIR:

Hostile:
In process of conducting a hostile act

Suspect:
Contact that has previously committed a hostile act but is no longer on an attack profile
Contact that displays suspicious or potentially threatening behaviour
Any Yemen military aircraft

Neutral
Any non-military aircraft acting in a non-threatening manner
Military aircraft belonging to a neutral nation

Assumed Friend
Single contact that correlates within 5 degrees of an ESM bearing of a friendly electronic emission.
Contact on same bearing as another showing mode 4

Friend
Visually sighted and recognized as a friendly military aircraft
Contact displaying a valid mode 4 IFF reply and is the only contact on that bearing

Unknown
Evaluated track that cannot be identified

SURFACE:

Hostile
In process of conducting a hostile act

Suspect
Contact that has previously committed a hostile act but is no longer on an attack profile

Contact that displays suspicious or potentially threatening behaviour
Any Yemen military vessel

Neutral
Any non-military vessel acting in a non-threatening manner
Military vessel belonging to a neutral nation

Friend
Contact correlates with friendly ESM
Contact visually identified as friend
Contact with proper mode 4 response

Unknown
Evaluated track that cannot be identified

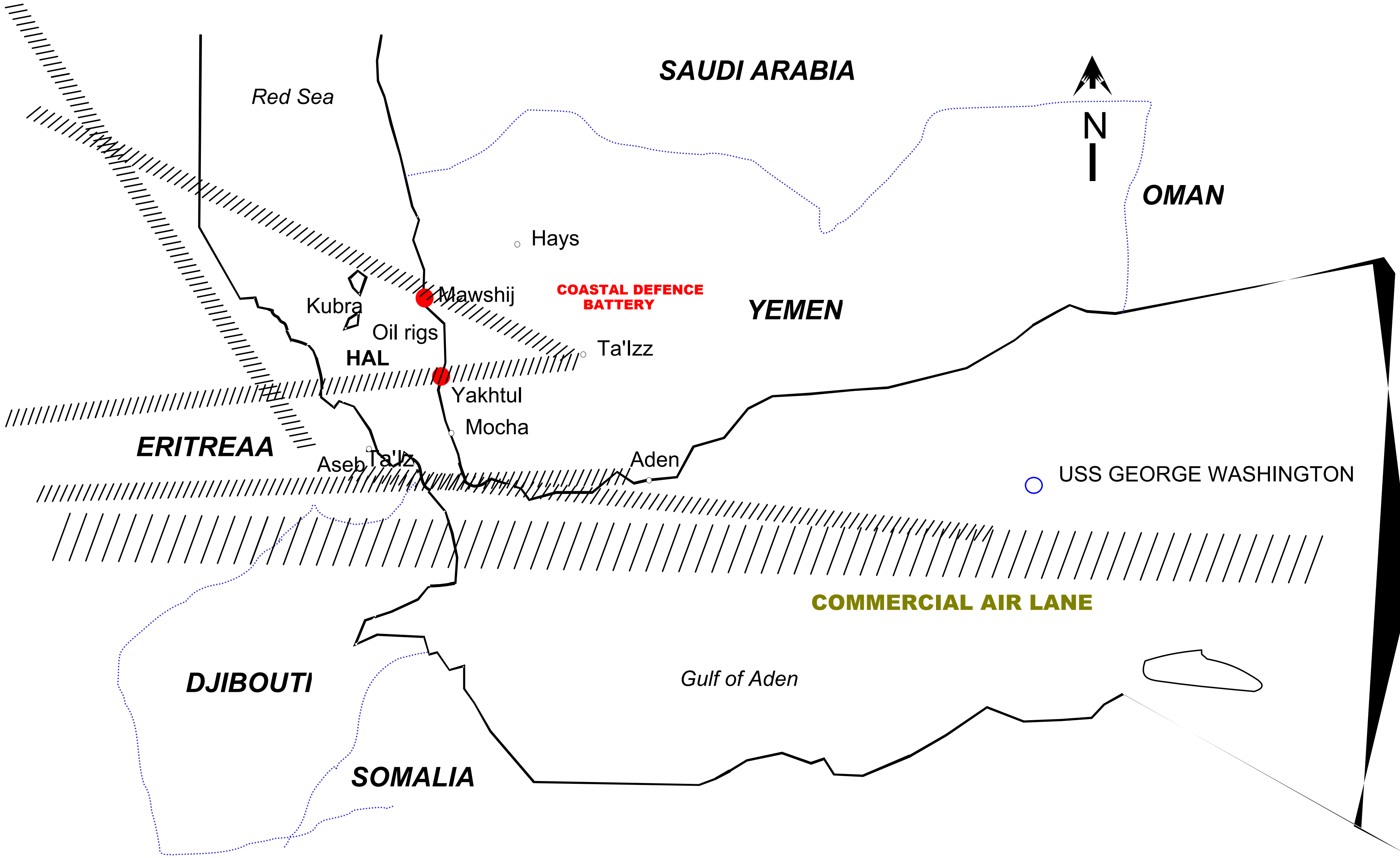
Yemen Order of Battle

Platform	Weapons	Emitters
Surface:		
3 X Tarantul	4 X SSN-2C 1 X 76mm 2 X 30mm CIWS	Search - Plank Shave Fire Control - Bass Tilt
2 X Coastal Missile Batteries	8 X SSN-2C each	Puff Ball
AIR:		
12 X Mirage	Exocet	Cyrano
22 X Fitter (FBA)	2 X 500lb bombs	High Fix
8 X Fishbed (Intercept)	Air to air	Jay Bird

Eritrea Order of Battle

Platform	Weapons	Emitters
AIR:		
22 X Fitter (FBA)	2 X 500lb bombs	High Fix

THIS PAGE INTENTIONALLY LEFT BLANK


SAUDI ARABIA
Red Sea
OMAN
Hays
Mawshij
COASTAL DEFENCE
BATTERY
YEMEN
Kubra
Oil rigs
HAL
Ta'Izz
Yakhtul
Mocha
ERITREAA
Aseb
Ta'Iz
Aden
USS GEORGE WASHINGTON
COMMERCIAL AIR LANE
DJIBOUTI
Gulf of Aden
SOMALIA
N

DOCUMENT CONTROL DATA SHEET

1a. PERFORMING AGENCY	2. SECURITY CLASSIFICATION
Humansystems, Inc. 111 Farquhar St., 2nd floor, Guelph, ON N1H 3N4	UNCLASSIFIED -
1b. PUBLISHING AGENCY DRDC Toronto	

3. TITLE

(U) Development of measures of performance for evaluating the comdat technology demonstrator: Potential use of the naval combat operator trainer (NCOT) for data collection

4. AUTHORS

M.L. Matthews R.D.G. Webb A.R. Keeble T. Lamoureux

5. DATE OF PUBLICATION	6. NO. OF PAGES
January 22 , 2004	121

7. DESCRIPTIVE NOTES

8. SPONSORING/MONITORING/CONTRACTING/TASKING AGENCY
Sponsoring Agency:
Monitoring Agency:
Contracting Agency : DRDC Toronto
Tasking Agency:

9. ORIGINATORS DOCUMENT NO.	10. CONTRACT GRANT AND/OR PROJECT NO.	11. OTHER DOCUMENT NOS.
Contract Report CR 2004-009	11bg01	

12. DOCUMENT RELEASABILITY

Unlimited distribution

13. DOCUMENT ANNOUNCEMENT

Unlimited announcement

14. ABSTRACT

(U) This report compiles the information contained in four other reports and Technical Memoranda pertaining to the development of Measures of Performance (MOP) for Multi-Sensor Data Fusion technology, and evaluates logistics and the utility of the Naval Combat Operator Trainer (NCOT) to gather these MOPs. Each report/Technical Memorandum is preceded by a summary of its content. This report recounts the development of specific MOPs, the conduct of a proof-of-concept trial at NCOT (with attendant identification of potential system improvements), the outcome of discussions with the contractor responsible for the maintenance and operation of NCOT regarding enhancements, and the investigation of one potential commercial-of-the-shelf solution for some of the improvements described. As a result of identifying these potential improvements, it was recommended that team records created during Navy team training in Operations Room Team Trainer (ORTT) be considered as a better source for deriving MOPs.

(U) Le présent rapport contient une compilation de l'information contenue dans quatre autres rapports et documents techniques portant sur l'élaboration de mesures de la performance pour la technologie de fusion de données de multicapteurs (MSDF), ainsi qu'une évaluation de la logistique et de l'utilité du simulateur d'opérateur d'équipement de combat naval (NCOT) en vue de la préparation de ces mesures de la performance. Chaque rapport/document technique est précédé d'un résumé de son contenu. Le présent rapport récapitule l'élaboration de mesures particulières de la performance, la conduite d'un essai de validation de concept au NCOT (l'opérateur identifiant les améliorations possibles à apporter au système), les résultats de discussions menées avec l'entrepreneur chargé de l'entretien et du fonctionnement du NCOT en ce qui concerne les améliorations et l'étude d'une solution commerciale courante possible pour certaines des améliorations décrites. Suite à la détermination des améliorations possibles, il a été recommandé que les dossiers créés par les équipes durant l'instruction d'équipes de la Marine au moyen du simulateur des équipes de salle des opérations (ORTT) soient considérés comme de meilleures sources pour l'établissement des mesures de la performance.

15. KEYWORDS, DESCRIPTORS or IDENTIFIERS

(U) naval command and control; evaluation; measures of performance